Trajectory Prediction Accuracy Report: User Request Evaluation Tool (URET)/ Center-TRACON Automation System (CTAS)

Mike M. Paglione
Dr. Hollis F. Ryan
Robert D. Oaks
J. Scott Summerill
Mary Lee Cale

May 1999

DOT/FAA/CT-TN99/10

U. S. DEPARTMENT OF TRANSPORTATION
Federal Aviation Administration
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

NOTICE

Technical Report Documentation Page

1. **Report No.**	2. **Government Accession No.**	3. **Recipient's Catalog No.**
DOT/FAA/CT-TN99/10		

4. Title and Subtitle
Trajectory Prediction Accuracy Report: User Request Evaluation Tool (URET)/Center-TRACON Automation System (CTAS)

5. Report Date
May 1999

6. Performing Organization Code
ACT-250

7. Author(s)
Mike Paglione, Dr. Hollis Ryan, Robert Oaks, J. Scott Summerill, Mary Lee Cale

8. Performing Organization Report No.
DOT/FAA/CT-TN99/10

9. Performing Organization Name and Address
U. S. Department of Transportation
FAA William J. Hughes Technical Center
ATC Engineering and Test Division
ATM Engineering, Research and Evaluation Branch
Atlantic City International Airport, NJ 08405

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
U. S. Department of Transportation
Free Flight Phase 1 Program Office
Conflict Probe/URET Product Team
Washington, D. C. 20590

13. Type of Report and Period Covered

14. Sponsoring Agency Code
AOZ-200

15. Supplementary Notes
The authors identified above represent the following organizations: Mike Paglione and Mary Lee Cale, ACT-250; Dr. Hollis Ryan and J. Scott Summerill, Signal Corporation; Robert Oaks, Crown Communications.

15. Abstract

This report presents the results of an independent analysis of the accuracy of the trajectory modelers implemented in the User Request Evaluation Tool (URET) and Center-TRACON Automation System (CTAS) prototypes. These results are based on the completion of the first phase of a planned two phased effort. As originally envisioned, efforts during Phase 1 would develop a generic methodology to measure the trajectory prediction accuracy of any decision support tool (DST), which would be validated by applying it to CTAS and URET based on their currently adapted sites. In Phase 2, the methodology would be applied to URET and CTAS adapted to a common site and supplied with the same scenario. As such, the results from Phase 2 would have provided a common set of results based on the same site and scenario, allowing a comparison of the two trajectory modelers to be made, in support of research into the performance requirements for a common en route trajectory model. Due to funding cuts, this task was curtailed to the completion of Phase 1. The results from this phase do provide the FAA with an independent set of scenario-based trajectory accuracy statistics for each DST, but they cannot be used to compare the two DSTs due to the confounding site-specific factors.

A methodology was developed and CTAS and URET were measured based on one scenario each from their currently adapted sites (Fort Worth and Indianapolis, respectively). The Phase 1 study measured the spatial error between trajectory predictions versus the Host Computer System (HCS) track position reports, which were assumed to be the ground truth location of the aircraft. The spatial error consisted of horizontal and vertical errors. The horizontal error was further partitioned into two geometric components, lateral and longitudinal errors, representing the cross track and along track prediction errors. The focus of the analysis was on the overall trajectory accuracy of each DST, not on individual errors. A statistical analysis was performed on the overall accuracy of each modeler and the spatial errors have been summarized with descriptive statistics in the horizontal, lateral, longitudinal, and vertical dimension as a function of look ahead time. Inferential statistics were performed to determine whether specific factors (e.g., look ahead time, flight type, horizontal phase of flight and vertical phase of flight) had a significant effect on these performance statistics.

While the Phase 1 analysis cannot be used to compare the URET and CTAS trajectory modelers, the results do provide the FAA with an independent scenario-based set of trajectory accuracy measurements for each DST. In addition, a generic methodology has been developed that can be used to determine the performance requirements for a common en route trajectory model.

16. Key Words

URET	trajectory
CTAS	longitudinal error
DST	lateral error
vertical error	HCS

17. Distribution Statement
This document is on file at the
FAA William J. Hughes Technical Center Library
Atlantic City International Airport, New Jersey 08405

19. **Security Classif. (of this report)**	20. **Security Classif. (of this page)**	21. **No. of Pages**	22. **Price**
Unclassified	Unclassified	556	

ERRATA

Report No. DOT/FAA/CT-TN99/10

TRAJECTORY PREDICTION ACCURACY REPORT:
USER REQUEST EVALUATION TOOL (URET)/
CENTER-TRACON AUTOMATION SYSTEM (CTAS)

May 1999

TECNICAL NOTE

Prepared for
U. S. DEPARTMENT OF TRANSPORTATION
Federal Aviation Administration
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

REPLACE PAGE 65

Released August 1999

Acknowledgements

Efforts upon which the contents of this report are based were planned and conducted by the FAA ATM Engineering, Research and Evaluation Branch (ACT-250) Conflict Probe Algorithm Assessment Team. These team members include Mary Lee Cale and Mike Paglione, ACT-250; Dr. Hollis F. Ryan and J. Scott Summerill, Signal Corporation; and Robert Oaks, Crown Communications, Inc. Mrs. Cale and Mr. Paglione were the FAA co-leads for this effort. Dr. Ryan developed the data reduction and analysis tools for aircraft track positions and traced many flights in detail, analyzing the causes of the trajectory prediction errors. Mr. Oaks designed and developed the generic trajectory sampling software and trajectory parsing software for the CTAS predictions. Mr. Summerill developed the phase of flight software and flight plan parsing tools, and performed much of the statistical analysis presented in this report.

The overall effort required to successfully accomplish this task was supported by several other members of ACT-250. In particular, Jeff Jackson of Signal Corporation supported the Oracle database system, which is a major part of the infrastructure upon which the data reduction and analysis tools reside. As system administrator in the Traffic Flow Management Laboratory, Craig Stewart of Signal Corporation provided continuous support and attention to the team's special needs, which is greatly appreciated. Richard Van Suetendael conducted a thorough technical review of the document and provided many insightful comments. Finally, special thanks go to Joan Feuerstein for her extensive help in preparing the document; without her, the timely completion of this report would not have been possible.

The authors appreciate the efforts of the Trajectory Working Group, of the FAA/NASA Interagency Air Traffic Management Integrated Product Team (IAIPT) En Route Area Work Team, who provided sound technical input and guidance during this study. Trajectory Working Group members include: Banavar Sridhar, Steven Green, and David McNally, NASA Ames Research Center; Daniel Kirk, MITRE/CAASD; and Herman Vandevenne and Steven Campbell at MIT Lincoln Laboratory. In particular, Dr.'s Green, Vandevenne, and Kirk provided detailed comments which have been incorporated into the final report.

Finally, the support and cooperation of both NASA Ames and MITRE/CAASD was invaluable in accomplishing this effort. Thanks to Karl Bilimoria at NASA Ames for his original suggestion on using his trajectory capturing technique in CTAS, and to David McNally for supplying the CTAS Fort Worth traffic data. Special thanks to Michelle Eshow at NASA Ames for her help in obtaining and modifying the Daisy View version of CTAS used in this study. Thanks to Maureen Tucker and Keith Lewis at MITRE/CAASD for their URET support and for supplying the URET Indianapolis traffic data.

Table of Contents

List of Figures

List of Tables

Executive Summary

In the spring of 1998, the FAA Air Traffic Management (ATM) Engineering, Research and Evaluation Branch (ACT-250) was tasked by the En Route Area Work Team lead (at the time, AUA-540), of the Interagency Air Traffic Management Integrated Product Team (IAIPT), to conduct an independent assessment of the technical accuracy of the User Request Evaluation Tool (URET) and Center TRACON Automation System (CTAS) aircraft trajectory modeling algorithms. This study was initiated under IAIPT Joint Research Project Description (JRPD) 57 in support of research into the performance requirements for a common en route trajectory model. The task was partitioned into two parts. In Phase 1, a generic methodology to measure trajectory prediction accuracy would be developed and validated by applying it to CTAS and URET at their currently adapted sites. For Phase 2 the same methodology would be applied to CTAS and URET adapted to a common site and supplied with the same scenario. Due to funding limitations in FY99, this task was curtailed to the completion of only Phase 1, which is documented in this report. As such, it provides the FAA with an independent scenario based analysis of URET and CTAS trajectory prediction accuracy but these results can not be used to compare the two modelers due to the confounding site-specific factors.

A generic methodology was developed to analyze any Decision Support Tool's (DSTs) trajectory modeling. This methodology took the point of view of an air traffic controller using the DST. That is, a Controller viewing the aircraft predicted position data on the graphical user interface of the DST would wonder how accurate the predictions were into the future, e.g., 5 minutes, 10 minutes, 20 minutes, and beyond. The Controller is not necessarily interested in the interior workings of the tool, e.g., how recently the tool made its currently valid predictions, but rather how accurate the prediction is now, and into the future. Built upon this conceptual point of view of the user, a sampling process was used to obtain the measurement data. At selected times the actual position of the aircraft was obtained from the HCS radar track data and was compared with the position of the aircraft predicted by the tool.

The results presented are based on field data collected at Fort Worth Air Route Traffic Control Center (ARTCC) in January 1999 for CTAS and in Indianapolis Air Route Traffic Control Center (ARTCC) in February 1998 for URET. Both scenarios were approximately 7 to 7.5 hours in duration and provided about 2500 flights for analysis. The analysis was performed on approximately 17,000 URET trajectories and 32,000 CTAS trajectories. The main focus of the analysis was on the overall trajectory accuracy of each DST. The spatial errors have been summarized with descriptive statistics in the horizontal, lateral, longitudinal, and vertical dimensions as a function of look ahead time. Inferential statistics were performed to determine whether specific factors (i.e., look ahead time, flight type, horizontal phase of flight, and vertical phase of flight) had a significant effect on these performance statistics. For URET, the sample means of the horizontal error as a function of look ahead time range from 1.2 to 10.2 nautical miles for zero to 30 minutes look ahead time. The sample standard deviations range from 1.1 to 10.9 nautical miles. For CTAS, the sample means of horizontal error as a function of look ahead time range from 0.3 to 10.9 nautical miles for 0 to 30 minutes look ahead time. The sample standard deviations range from 0.9 to 11.2 nautical miles. For both URET and CTAS, the average and standard deviation of the horizontal error increased as look ahead time increased. In other words, the horizontal uncertainty of the trajectory predictions analyzed in this study increased by about 10 nautical miles on average as look ahead increased from zero to 30 minutes into the future.

While the Phase 1 analysis cannot be used to compare the URET and CTAS trajectory modelers, the results do provide the FAA with an independent scenario based set of trajectory accuracy measurements for each DST. All of the data from this study is stored in a large set of Oracle database tables in the WJHTC TFM Laboratory. This data can be made available to other members of the FAA community who may wish to analyze other factors, or answer other questions of interest, related to the trajectory prediction accuracy of URET and CTAS upon formal request to ACT-250. In addition, a generic methodology has been developed for the performance measurement of a common trajectory model. In FY99, this methodology and the parsing tools developed in this study will be applied to the development of DSR Workload Scenarios to be used for URET CCLD accuracy testing. With the planned adaptation of URET and CTAS to a common site (tentatively scheduled to occur in 2001) and anticipated funding availability in FY01, ACT-250 hopes to resume work on the proposed Phase 2 study to further address the FAA's efforts to determine the feasibility of a common en route trajectory model.

1. Introduction

1.1 Purpose

This report presents the results of an independent analysis of the accuracy of the aircraft trajectory modelers implemented in the User Request Evaluation Tool (URET) and the Center-TRACON Automation System (CTAS) prototypes. This study was conducted by the Air Traffic Management (ATM) Engineering, Research and Evaluation Branch (ACT-250) at the FAA William J. Hughes Technical Center (WJHTC). Quantitative measures of the trajectory accuracy of URET and CTAS are presented in terms of the following metrics: horizontal error (longitudinal error and lateral error) and vertical error. These results are based on analyses of field data obtained from the Indianapolis and Fort Worth Air Route Traffic Control Centers (ARTCCs) where the URET and CTAS prototypes, respectively, are currently implemented; as such, while this report provides useful information on the accuracy of the individual tools, the results cannot be used to compare the performance of the trajectory modelers.

1.2 Background

To achieve the goals of Free Flight, broad categories of advances in ground and airborne automation are required. The FAA has sponsored the development of two ground based ATM decision support tools (DSTs) to support the en route and arrival air traffic controllers. URET, developed by MITRE/CAASD, facilitates the controller's management of en route air traffic by identifying potential air traffic conflicts. CTAS, developed by NASA Ames Research Center, supports the controller in the development of arrival sequencing plans and the assignment of aircraft to runways to optimize airport capacity. A fundamental component of both URET and CTAS is the trajectory modeler, upon which the functionality provided by these tools is based. For example, URET uses its predicted trajectories to predict conflicts; CTAS uses its predicted trajectories to calculate meter fix crossing times. Thus, the trajectory accuracy, or the deviation between the predicted trajectory and the actual path of the aircraft, has a direct effect on the overall accuracy of the tool.

The prediction accuracy of URET and CTAS is a critical issue to be addressed in planning for Free Flight Phase 1 (FFP1) and the future integration of these tools. NASA Ames Research Center and MITRE/CAASD have each created and applied performance metrics for their specific tools (Bilimoria, 1998; Brudnicki et. al., 1998). The ATM Engineering, Research and Evaluation Branch (ACT-250) at the FAA WJHTC has defined a generic set of metrics that highlight the performance of any decision support tool: trajectory accuracy, conflict prediction accuracy, prediction stability and conflict notification timeliness (WJHTC/ACT-250, 1997 and Cale et al., December 1998). Since these metrics are independent of a particular system's design choices, they provide common measures to evaluate the performance of different systems. In early 1998, ACT-250 applied the conflict prediction accuracy metrics to URET (Cale et al., April 1998). Following the completion of the URET conflict prediction accuracy assessment, ACT-250 was tasked by the Interagency ATM Integrated Product Team (IAIPT) En Route Area Work Team lead (at that time, AUA-540) to conduct an independent assessment of the technical accuracy of the CTAS and URET trajectory modeling algorithms. This report focuses on the initial application of the trajectory accuracy metrics to URET and CTAS.

1.3 Scope

ACT-250's original plan for the trajectory accuracy study called for a two-phased effort. During the first phase, the necessary data reduction and analysis tools would be developed and validated by applying them to URET and CTAS based on the ARTCCs to which these DSTs were currently adapted (i.e., Indianapolis and Fort Worth). Phase Two then called for both systems to be adapted to a common ARTCC, with the trajectory accuracy study conducted based on this common data and a report issued. Toward the end of Phase One, funding was cut for ACT-250's IAIPT tasks for FY99 and ACT-250's focus shifted to the development of scenarios to be used for the FFP1 URET Core Capability Limited Deployment (CCLD) accuracy testing. Since the initial trajectory study was almost completed and many of the tools being developed were required by the scenario development task, it was decided to complete this study and provide a report even thought the results are limited to the Phase One effort. Therefore, while the results presented provide an estimation of the accuracy of the individual tools' trajectory modelers, this data can not be used to compare the two modelers because it is based on information from two different centers at different time periods with different weather characteristics.

1.4 Document Organization

This report is organized into five sections and three appendices. Section 2 provides a detailed description of the methodology employed to conduct the trajectory accuracy study. Sections 3 and 4 describe the scenarios, and observations and results for the URET and CTAS studies, respectively, and Section 5 provides a summary of the study. Document references and a list of acronyms are also provided. In addition, three appendices are provided: detailed descriptions of the data analyzed for each tool are provided in Appendix A, standard deviation statistical plots of results are provided in Appendix B, and additional flight observation examples are provided in Appendix C.

2. Trajectory Accuracy Study Methodology

The WJHTC ATM Engineering, Research and Evaluation Branch (ACT-250) has been involved in the development and application of metrics to assess various aspects of decision support tools since early 1997 (WJHTC/ACT-250, 1997; WJHTC/ACT-250, 1998; Cale et. al, April 1998; Cale et. al, December 1998). The fundamental characteristic of these metrics is their independence from any particular DST's design choices, thus providing common measures to evaluate the performance of different systems. The approach employed for this study used field data recorded at two of the ARTCCs where the URET and CTAS prototypes are currently implemented.

The effective estimation of the trajectory accuracy metrics required considerable data to be collected and analyzed. A generic set of data reduction and analysis tools was developed, building upon ACT-250's Traffic Flow Management (TFM) Laboratory's Oracle database system and tools previously developed for the URET Conflict Prediction Accuracy Study (Cale et. al., April 1998). This section describes these generic techniques as they were used in this trajectory study, and provides information on the definitions used throughout the study, the sources of data and the data processing methodology, the data processing reports that were generated, and the analysis performed. Sections 3.2 and 4.2 contain observations for URET and CTAS, respectively, that demonstrate the application of this methodology.

2.1 Overview

Three major process areas comprise the Trajectory Accuracy Study methodology (shown in Figure 2.1-1):

1. **Field Data Parsing** - The recorded field data, which may be provided in different formats, is parsed to extract the flight plan data, the track data, and the trajectory data into a common format. The details of this DST-specific parsing are provided in Sections 3.1 and 4.1.
2. **Flight Plan and Track Data Processing -** The software in this process area filters and characterizes the track data, placing the results in tables in the TFM Laboratory Oracle database. Details on this processing are provided in Section 2.4.
3. **Trajectory Data Processing and Trajectory Report Generation -** During these processes, the trajectory data is sampled and compared with the track data, the metrics are calculated and placed into tables in the TFM laboratory Oracle database, and reports are generated. Trajectory data sampling is necessary due to the differences in trajectory creation methods employed by URET and CTAS (i.e., CTAS computes a new trajectory every 12 seconds for every track update, while URET's trajectory creation is mainly event driven); on average, 10-12 times more trajectories were created for CTAS than for URET. Because of this, a sampling technique was designed to create equivalent sets of trajectory data for analysis. Details on this processing are provided in Section 2.5.

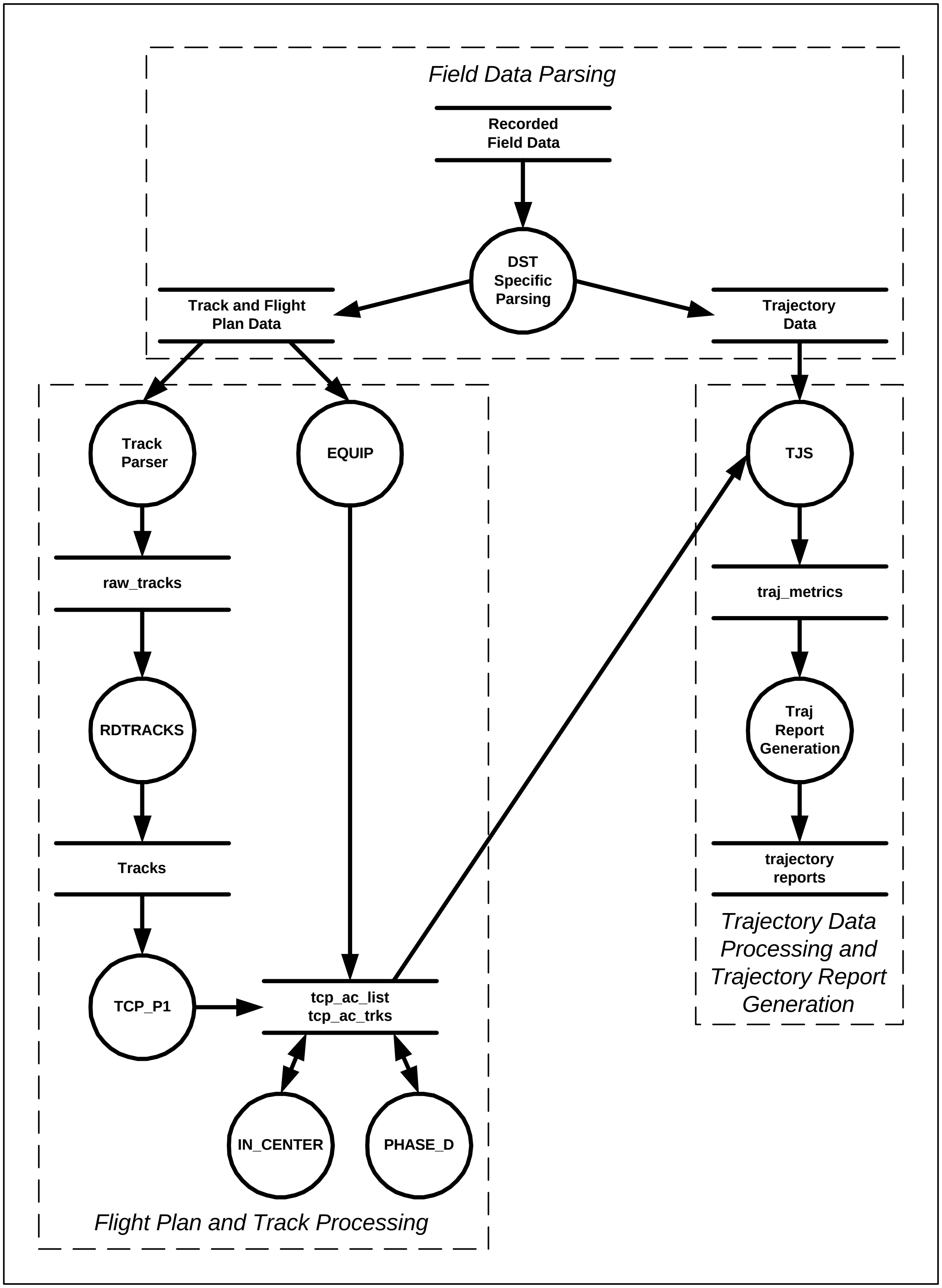

Figure 2.1-1: Trajectory Accuracy Study Methodology Overview

2.2 Definitions

This section defines the basic terms used throughout this report. These are grouped into three categories: data definitions, metrics definitions, and factor definitions.

2.2.1 Data Definitions

Three types of data were used as input to this study: flight plan, track, and trajectory data.

2.2.1.1 Flight Plan Data

A flight plan consists of time stamped records containing information about the aircraft's flight, including: aircraft identification (ACID), computer identification number (CID), aircraft type, coordination fix, coordination time, and intended route of flight. For both URET and CTAS, the flight plan data for this study was recovered from flight plan and flight plan amendment messages output from the ARTCC Host Computer System (HCS) and recorded by the URET or CTAS interface software.

2.2.1.2 Track Data

Track data represents the position of an aircraft as reported by the ARTCC HCS. An aircraft's track is represented by a sequence of four-dimensional data points, with each data point consisting of three spatial coordinates (denoted X_i, Y_i, and Z_i) and their associated time (denoted T_i,), where i refers to a particular data point. For both URET and CTAS, the track data for this study was recovered from track messages output from the ARTCC HCS and recorded by the URET or CTAS interface software.

2.2.1.3 Trajectory Data

Trajectory data represents the position of an aircraft as predicted by the DST into the future. A trajectory is a sequence of four-dimensional data points, with each data point consisting of three spatial coordinates (denoted X_i, Y_i, and Z_i) and their associated time (denoted T_i,), where i refers to a particular data point. The trajectory data for this study was directly captured from the URET and CTAS trajectory modelers.

2.2.2 Metrics Definitions

Trajectory accuracy can be measured as the spatial difference between the predicted path of the aircraft determined by the DST and the aircraft's actual path. This difference is the slant range distance between the predicted trajectory position and the actual track position at a common time. A perfect prediction would have a slant range of zero.

For this study, trajectory accuracy was measured as the difference between the URET or CTAS predicted trajectory and the tracked position reports received from the ARTCC HCS. This slant range distance was decomposed into three orthogonal components: longitudinal error and lateral error in the horizontal plane, and vertical error perpendicular to the horizontal plane. Both the longitudinal and lateral errors are also orthogonal components of the horizontal error. The horizontal error is the slant range's projection onto the horizontal plane. These errors are actually vectors, however statistical analysis was performed only on their scalar lengths and a sign convention was used for direction, where appropriate. The details for estimating these metrics are presented in Section 2.5.1.2.

2.2.2.1 Longitudinal Error

The longitudinal error represents the along track distance difference between a track and its trajectory. This error, depicted in Figure 2.2-1, lies in the horizontal plane defined by a track point and two consecutive trajectory points. As seen in Figure 2.2-1, a positive longitudinal error indicates that at a corresponding point in time the aircraft is ahead of where the trajectory predicted it would be.

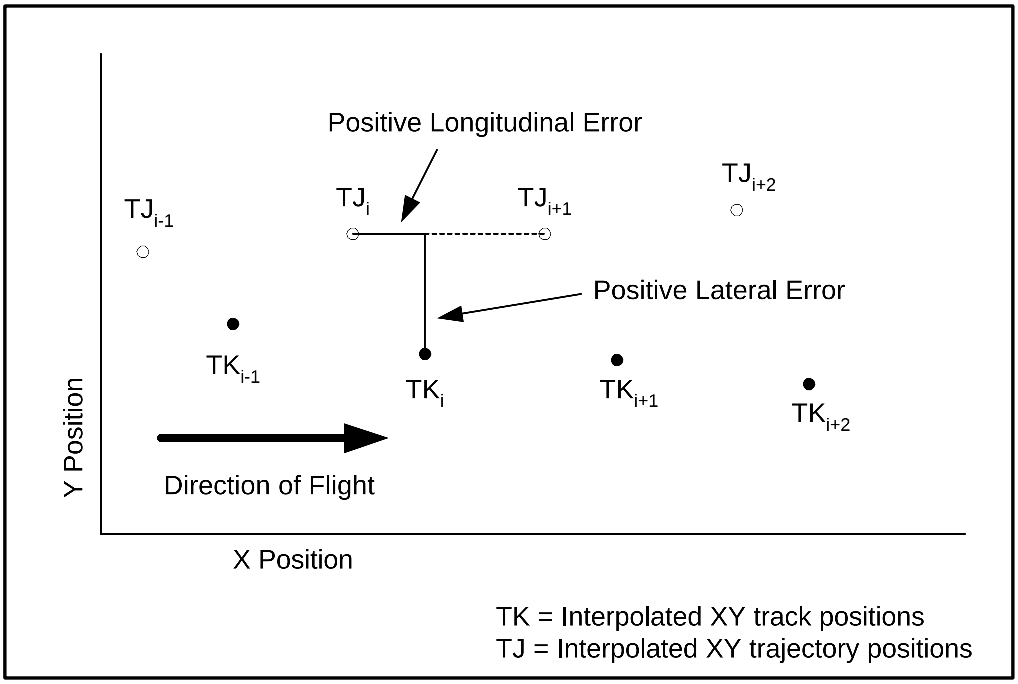

Figure 2.2-1: Longitudinal and Lateral Errors

2.2.2.2 Lateral Error

The lateral error represents the side to side, or cross track, difference between a track and its trajectory. This error, also represented in Figure 2.2-1, lies in the horizontal plane defined by a track point and two consecutive trajectory points. A positive lateral error indicates that the aircraft is to the right of the predicted trajectory at a corresponding point in time.

2.2.2.3 Vertical Error

The vertical error represents the difference between the tracked altitude and the predicted altitude. This error, depicted in Figure 2.2-2, lies perpendicular to the horizontal plane. A positive vertical error indicates that at a corresponding point in time the aircraft is above where the trajectory predicted it would be.

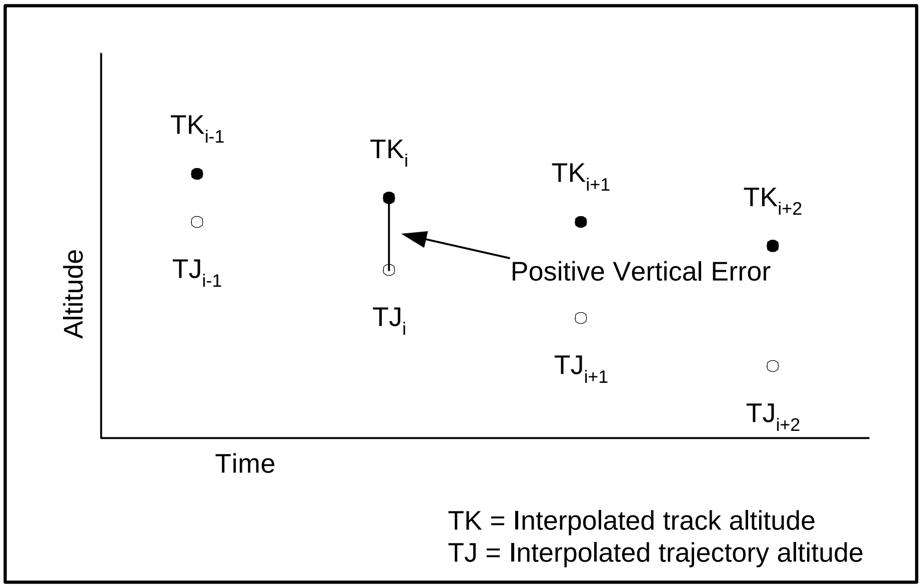

Figure 2.2-2: Vertical Error

2.2.3 Factors Definitions

Various factors that have the potential of affecting the accuracy of a trajectory modeler were examined during this study. These factors, which include trajectory build time, early trajectory, look ahead time, phase of flight, flight type and aircraft type, are defined in the following sections.

2.2.3.1 Trajectory Build Time

During the life of an aircraft track, a trajectory modeler computes numerous trajectories, each with an associated build time. Since, the trajectory accuracy metrics were computed at a number of sample times along an aircraft track it was necessary to establish criteria for selecting which trajectory to use in these computations. The trajectory selected for a specific sampling time along an aircraft track was the trajectory with the most recent build time, not exceeding the sample time. The determination of this factor is described in Section 2.5.1.1.

2.2.3.2 Early Trajectory

Depending on the method employed for creating trajectories (i.e., upon receipt of every track point or event driven), it is possible for a trajectory to be computed before the start of the track data. For this study, these are identified as “early trajectories”. These trajectories are built strictly with the flight plan without HCS track information. The determination of this factor is described in Section 2.5.2.

2.2.3.3 Look Ahead Time

Associated with the error measures for a pair of points is a look ahead time. This look ahead time is the difference between the time point at which the metrics are computed for a sampled trajectory/track position and a base time. The base time represents the first calculation of the metrics taken among a sequence. The sequence starts by taking the current track point and a time coincident trajectory point off the currently available trajectory. The first point is the base time and then every parameter number of seconds, or look ahead time, into the future the metrics are calculated on this same trajectory. The sequence iterates again every parameter number of seconds based on the sampling methodology defined in Section 2.5.1.1.

It is important to note that the look ahead time is based on the start of each sampling interval and is not directly related to the age of the trajectory as defined in other studies. For example, MITRE/CAASD defines look ahead time to be the difference between the trajectory build time and the time into the future a metric is calculated along that trajectory (Brudnicki, August 1995). In the ACT-250 study definition, a look ahead time of zero may be calculated on a trajectory that has an age of more than zero. The determination of this factor is described in Section 2.5.

2.2.3.4 Phase of Flight

In the horizontal plane an aircraft can be considered to be either flying straight or turning. In the vertical plane an aircraft can be considered to be either flying level, ascending, or descending. The determination of these factors is described in Section 2.4.6.

2.2.3.5 Flight Type

With respect to an ARTCC, an aircraft can be considered to be:

- overflight - the aircraft track begins outside the center boundary, flies through the center, then ends outside the center boundary
- departure - the aircraft track begins at an airport within the center and ends outside the center boundary
- arrival - the aircraft track begins outside the center boundary and ends at an airport within the center
- internal - the aircraft track begins and ends at an airport within the center.

The details for estimating this factor are presented in Section 2.4.2.

2.2.3.6 Aircraft Type

The aircraft type is available as a part of an aircraft's flight plan message. For both DSTs, the aircraft type is an important factor in modeling the aircraft’s flight profile. The frequency of the top 20 aircraft types were reported for each data set used (see Sections 3.1 and 4.1), however an analysis of the effect of the aircraft type as a factor was left for future study.

2.3 Data Sources

The source of the flight plan and track data used for this study was recorded at the Indianapolis (ZID) and Fort Worth (ZFW) ARTCCs. Section 2.4 describes the generic techniques used to process this data. Specific data processes and procedures required for URET and CTAS are described in Section 3.1 and Section 4.1, respectively.

2.4 Flight Plan and Track Data Processing

Figure 2.4-1 provides a data flow diagram logically describing the data files and processes used to process the flight plan and track data. This processing was automated through a UNIX shell script that performed numerous functions through six primary processes: Track Parser, EQUIP, RDTRACKS, TCP_P1, IN_CENTER, and PHASE_D. These are further described in the following subsections.

2.4.1 Track Parser

The Track Parser process consists of a UNIX shell script and C++ programs that parse and sort the track data for input into the Oracle database table RAW_TRACKS.

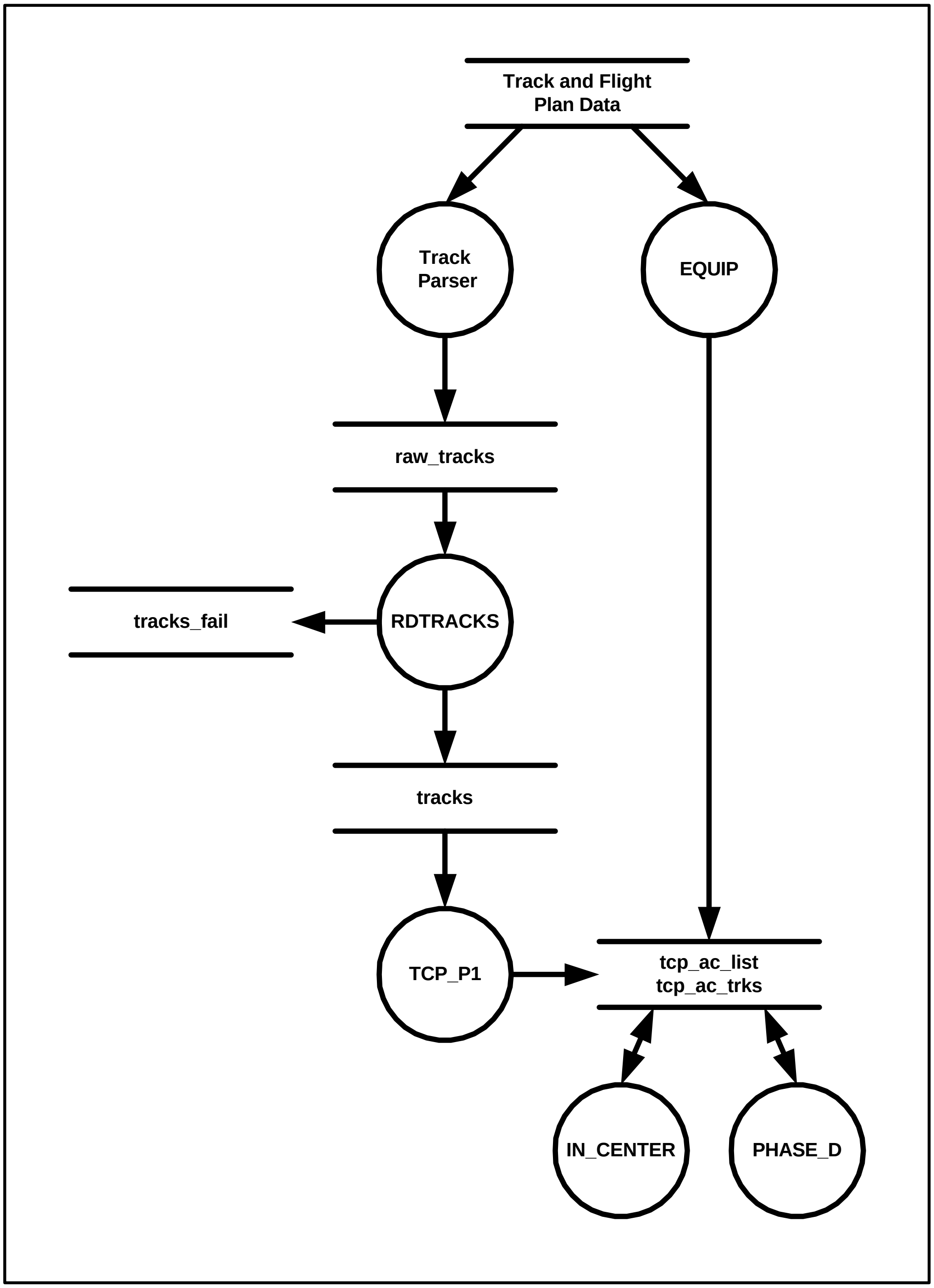

Figure 2.4-1: **Flight Plan and Track Data Processing**

2.4.2 EQUIP

EQUIP is a C program that extracts flight plan information and places it in the *tcp_ac_list* and *tcp_ac_trks* Oracle database tables. This information includes: the origin and destination airports, the flight type (arrival, departure, overflight, and internal), and the aircraft type and its equipage.

2.4.3 RDTRACKS

RDTRACKS is a C program that filters the RAW_TRACKS data to establish the "ground truth" tracks to which the trajectories are compared. RDTRACKS uses the URET and CTAS HCS tracks from their respective interface recorder files and produces files that are loaded into the TFM laboratory Oracle database to create the *tracks* and *tracks_fail* tables (described in WJHTC/ACT-250, 1999). The specific functions performed by RDTRACKS are described in the following subsections.

2.4.3.1 Correction of HCS Radar Track Position Reports

The radar track data supplied to the tools by the HCS contained inaccuracies and needed to be corrected before the error measurements could be made against the DST's trajectory predictions. For example, the following faults were found in the HCS track data:

- **Missing Track Reports -** Nominally the HCS supplies a new track report every 12 seconds. However, there were situations where the HCS omitted track reports, creating a gap in the position data (occasionally five or 10 minutes long). Short gaps (time gaps of less than two minutes) were patched by linear interpolation in all dimensions. Long gaps (time gaps of more than two minutes) were not patched and no accuracy measurements were made for these sections of the aircraft flight path.
- **Stationary Track Reports -** Frequently the HCS gave two or more successive track reports that had identical values for X, Y, and Z. That is, according to the HCS the aircraft had not moved (usually the HCS caught up with the next track report). This problem was fixed by linear interpolation.
- **Inconsistent Track Reports -** Because of its inertia, an aircraft is not able to make abrupt changes in velocity and position. Therefore, the distance traveled between position reports changes slowly. An abrupt change in track step size is not physically possible. A position report was considered to be inconsistent with the previous track report when an abrupt change occurred. Usually the position reports became consistent within a few track reports. Small amounts of inconsistent data were patched (i.e. less than two minutes), while large amounts (i.e. greater than two minutes) were not patched and measurements were not made during or beyond these gaps.
- **Jitter –** The position reports "bounce around" rather than following a smooth track as the aircraft is actually doing. This effect is noise or jitter on the position reports and is fairly small. It may be that the jitter exists in spite of the smoothing that the HCS does on the radar reports because of errors in the time data reported. As usual for real time processing systems, the data is not time stamped when it is collected. The time stamp is added later with reduced accuracy. For the statistical analyses performed in this study, the jitter was ignored. However, future studies may remove this additional source of error, using data smoothing techniques.

In addition to the track faults, there are differences in the methods by which URET and CTAS time stamp the track position reports from the HCS. RDTRACKS requires equally spaced track position reports, which URET supplies. CTAS track reports are not time stamped at equally spaced 12 second intervals but exactly as received downstream from the HCS interface.

Therefore, it was necessary to recover the HCS time values. This was done by rounding to the nearest whole second value and then these rounded values were rounded to the nearest integer multiple of 12 seconds. This was done in such a way as to minimize the total time adjustments for the entire track of the aircraft.

2.4.3.2 Track Processing Steps

The following processing was done to establish a good track history for an aircraft. If one value in a track report failed a test, the entire record was discarded. These tests did not ensure that a track report was accurate, but track reports that were clearly in error were excluded. If a track could not be initialized, the aircraft was not used in the study. At the start of each flight's track reports or following a large gap in time or spatial inconsistency, a flight's tracks are initialized. The initialization and continuous processing of the HCS track data is described below.

- **Prune Leading and Trailing Zeros -** Often the first one or two track reports for an aircraft had zero values for altitude. Similarly the last few records sometimes had zero altitude values. These reports were discarded.
- **Initialize Track -** The track was initialized by finding three good, contiguous track reports. A track report was considered good if it passed three tests:
  1. **Values Test -** The values test was used to catch gross errors in the aircraft position data. To pass the Values Test, Z had to be greater than zero and the absolute values of X and Y had to be less than 1000.
  2. **Delta Time Test -** To pass the Delta Time Test, the time of the track report had to be 12 seconds later than the time of the immediately preceding track report.
  3. **Fixed Delta Values Test -** To pass the Fixed Delta Values Test, the position of the aircraft in the horizontal (XY) plane must not have changed (in one 12 second step) by more than a maximum threshold value (3.0 nautical miles) nor less than a minimum threshold value (0.1 nautical miles). These threshold values correspond to aircraft speeds of 900 knots and 30 knots, respectively. In addition, the altitude of the aircraft could not have changed by more than a threshold value of 2000 feet, which corresponds to a climb or descent of 10,000 feet per minute. (Note that military aircraft were excluded from this study.)

After three good, contiguous track reports were found, the above three tests were repeated for each successive track report. Every record that passed all of the tests was passed unchanged to the next processing step in TCP_P1. If a report failed a test, an attempt was made, usually successfully, to fix the record by inserting new values obtained through interpolation between a previous good report and a later good report. There were two cases to handle: a time gap (missing data), and a bad data gap (one or more records were in error).

- **Time Gap Processing** - When a time gap in the data was found, a search was started for an acceptable next track report, starting with the current track report. Each successive track report was tested in turn. An acceptable next track report had to pass three tests: the Values Test described above, and the Variable Delta Values Test and the Maximum Time Gap Test, described below:
  1. **Variable Delta Values Test -** A prediction was made of where the aircraft would be if it maintained the same ground velocity as it had before the time gap. This predicted position was compared to the position reported by the candidate track report. The test was passed if the two positions were close enough to each other (three nautical miles). The average ground velocity was calculated using the last four position reports before the time gap.

2. **Maximum Time Gap Test -** The Maximum Time Gap Test determined if the time difference between the last good track report and the candidate next good report was less than or equal to two minutes. It was assumed that track data can be interpolated accurately for a time gap less than two minutes. This parameter setting of two minutes allowed up to nine successive position reports to be interpolated.

If a candidate track report failed either the Values Test or the Variable Delta Values Test, or both, the next track report was selected for testing. If the candidate track report passed the Values Test and Variable Delta Values Test, but failed the Maximum Time Gap Test, the track was re-initialized, whenever possible. If the track could not be re-initialized, it was terminated. If the search reached the end of the track data without finding a record which had passed all three tests, the track was terminated. If the candidate track report passed all three tests, it was output and used with the last good report to estimate, using linear interpolation, the missing track report positions in the time gap. The interpolation inserted track reports into the missing time slots and also replaced the track reports which failed the tests in the search for the next good report.

- **Bad Data Gap Processing -** A bad data gap was detected when a track report passed the Delta Time Test and the Values Test but failed the Fixed Delta Values Test. A search was then started to find the next good record. The search process was the same for a bad data gap as for a time gap. A search was started for an acceptable next track report, starting with the current track report. Each successive track report was tested in turn. An acceptable next track report had to pass three tests: the Values Test, the Variable Delta Values Test and the Maximum Time Gap Test, described above. When a candidate track report was found which passed all three tests, it was output and used with the last good report to estimate, using linear interpolation, the correct values of X, Y, and Z for the track report positions in the bad data gap. The interpolation inserted the corrected values into the track reports in the bad data gap. Then regular track processing was resumed. If a good next report could not be found, the track was terminated. If a next report passed the Values and the Variable Delta Values Tests, but failed the Maximum Time Gap Test, the track was re-initialized, if possible. If the track could not be re-initialized, it was terminated.

2.4.4 Track Conflict Probe

TCP_P1 is an Oracle Standard Query Language Plus (SQL/Plus) program that performs the interpolation of the track data. Although the HCS track reports normally are generated at 12-seconds intervals, for this study the track data was interpolated using a uniform 10-second time interval and synchronized with the hour.

An example of the relationship between recorded field data and interpolated aircraft tracks is shown in Figure 2.4-2. In this figure the **X**'s represent positional data generated by RDTRACKS at four time points. This data is specified in a time-of-day form and represents the aircraft's position at 16:25:13, 16:25:25, 16:25:37, 16:25:49, and 16:26:01. The **O**'s represent the interpolated positions with the time specified as the number of seconds elapsed since midnight. This interpolation was calculated using the MITRE/CAASD URET function CFP_POSIT (see Cale et. al., 1997, Section 3.1.9). This function uses a 2^{nd} order method in which the acceleration is assumed to be constant throughout the interpolation interval. The ground speeds are needed as input for the quadratic interpolation method; if they are not available this method degenerates to a linear interpolation method.

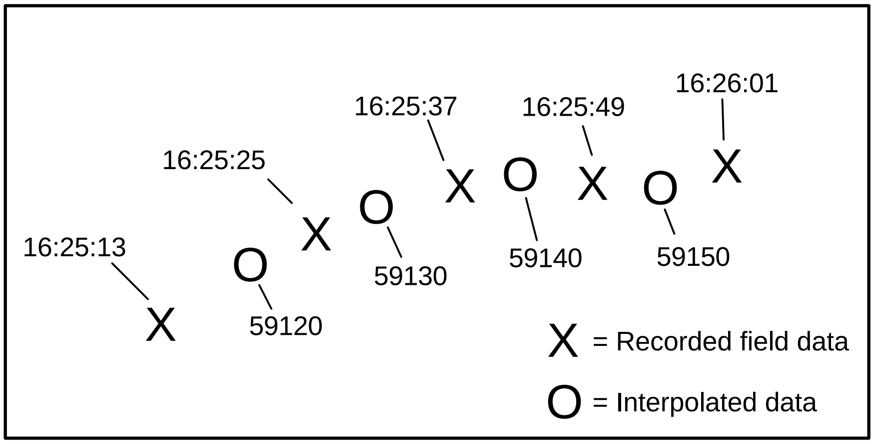

Figure 2.4-2: Interpolation of Recorded Aircraft Track Data

The following equations were used for quadratic interpolation where:

x is the desired interpolated X coordinate at time t,
x_1 is the value of x at time t_1,
x_2 is the value of x at time t_2,
y is the desired interpolated Y coordinate at time t,
y_1 is the value of y at time t_1,
y_2 is the value of y at time t_2, and $t_1 < t < t_2$.

In addition, for quadratic interpolation it was assumed that the acceleration was constant over the interpolation interval. The acceleration was then equal to the difference of the velocities at the start and end points of the interval divided by the length of the interval in time.

Let,

v_1 be the velocity of the aircraft at time t_1,
v_2 be the velocity of the aircraft at time t_2,
v_{1x} and v_{1y} be the X and Y components of the velocity v_1, and
v_{2x} and v_{2y} be the X and Y components of the velocity v_2.

Then the interpolated coordinate positions are

$$x = \frac{Ax_1 + Bx_2}{C} \qquad \textbf{Equation 2.4-1}$$

and

$$y = \frac{Dy_1 + Ey_2}{F} \qquad \textbf{Equation 2.4-2}$$

where

$$A = (v_{1x} - v_{2x})(t_2 - t)^2 + 2v_{2x}(t_2 - t_1)(t_2 - t) \qquad \textbf{Equation 2.4-3}$$

$$B = (v_{2x} - v_{1x})(t - t_1)^2 + 2v_{1x}(t_2 - t_1)(t - t_1) \qquad \textbf{Equation 2.4-4}$$

$$C \;=\; (v_{1x} + v_{2x})(t_2 - t_1)^2 \qquad \textbf{Equation 2.4-5}$$

$$D \;=\; (v_{1y} - v_{2y})(t_2 - t)^2 \;+\; 2v_{2y}(t_2 - t_1)(t_2 - t) \qquad \textbf{Equation 2.4-6}$$

$$E \;=\; (v_{2y} - v_{1y})(t - t_1)^2 \;+\; 2v_{1y}(t_2 - t_1)(t - t_1) \qquad \textbf{Equation 2.4-7}$$

$$F \;=\; (v_{1y} + v_{2y})(t_2 - t_1)^2 \qquad \textbf{Equation 2.4-8}$$

2.4.5 IN_CENTER

The IN_CENTER process determines if the interpolated track points fall within the center boundary. It uses an algorithm very similar to the MITRE/CAASD URET GM_REGN function (see Cale et. al., 1997, section 3.4.17) which determines if aircraft are within a protected or inhibited airspace. Since this study's application of this program was only interested in the end of an aircraft's track reports, all tracks were first flagged to be inside the center boundary. The algorithm was adapted to flag whether the track was outside the center boundary, starting from the end of the track reports and going backwards in time order. Processing was stopped for a flight's track as soon as it re-entered the center's airspace. For example, if an overflight had 100 interpolated track reports whose last 10 tracks were outside the center boundary (i.e. the 91st to 100th), this process determined each of the last 10 reports to be outside the Center boundary and the processing was terminated on the 90th track report when it was determined to be inside the Center.

The flag of inside or outside a center boundary, applied to the end of a flight's interpolated tracks, is utilized in the trajectory sampling process, since the trajectory prediction on tracks at the end of a flight outside the center are not processed for spatial prediction errors. This is an approximate method of excluding error calculations on the end portion of a flight transferring to another ARTCC and thus to another HCS and DST not included in the study.

2.4.6 PHASE_D

PHASE_D is a C program that determines the phase of flight of the aircraft in the horizontal and vertical directions, as discussed in the following subsections.

2.4.6.1 Horizontal Phase of Flight

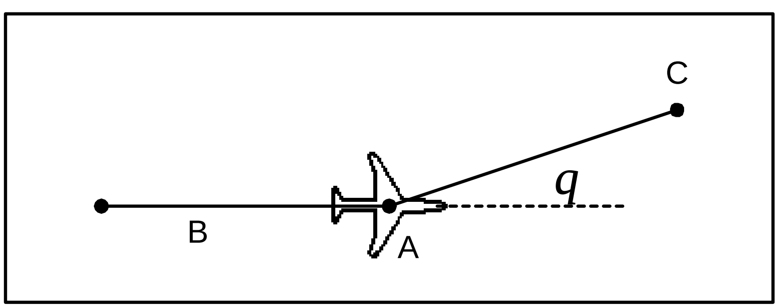

Figure 2.4-3: Horizontal Phase of Flight

The horizontal phase of flight for an aircraft, with respect to the ground, was defined as one of three states: straight, turning right, or turning left. The state was determined as follows: The point labeled A in Figure 2.4-3 represents the track point at which the aircraft's horizontal phase of flight is being determined. The point labeled B is a point along the interpolated track a parametric number of points (one point in this study) earlier in time than the point being examined. The point labeled C is a point along the track a parametric number of points (one point in this study) later in time than the point being examined. Then the vector V is defined as the normalized vector cross product of the vector from point B to A and the vector from point A to C, i.e.:

$$V = \frac{V_{BA} \times V_{AC}}{\|V_{BA}\| \|V_{AC}\|} \qquad \textbf{Equation 2.4-9}$$

where

V_{BA} is the vector defined by joining B to A

and

V_{AC} is the vector defined by joining A to C.

The magnitude of the vector V is the sine of the local change in bearing angle of the aircraft and can be used to determine the horizontal phase of flight, i.e., if the aircraft is flying straight this angle will be zero or close to zero. If the aircraft is turning the sine will not be close to zero and the sign of the sine of this angle will indicate whether the aircraft is turning left or right.

Since the vectors V_{BA} and V_{AC} are in the horizontal XY plane their vector cross product V is a vector perpendicular to the horizontal plane; i.e., coincident with the vertical or Z axis. In the NAS ARTCC coordinate system up is positive and down is negative. Therefore the sense of V is positive for a left turn and negative for a right turn. To determine whether the aircraft is flying straight or turning, the magnitude of V is compared to a threshold to minimize the effect of track position noise on the measurement.

Let the coordinates of the point A be x_a and y_a, the coordinates of the point B be x_b and v_b, and the coordinates of the point C be x_c and y_c. Then the components of the vectors V_{BA} and V_{AC} are:

$$V_{BA} = \begin{bmatrix} x_a - x_b \\ y_a - y_b \\ z_a - z_b \end{bmatrix} = \begin{bmatrix} v_{BAx} \\ v_{BAy} \\ v_{BAz} \end{bmatrix} \qquad \textbf{Equation 2.4-10}$$

$$V_{AC} = \begin{bmatrix} x_c - x_a \\ y_c - y_a \\ z_c - z_a \end{bmatrix} = \begin{bmatrix} v_{ACx} \\ v_{ACy} \\ v_{ACz} \end{bmatrix} \qquad \textbf{Equation 2.4-11}$$

Since the vectors are defined to be in the horizontal plane, the z components are all zero. The norms or magnitudes of the vectors are:

$$\|V_{BA}\| = \sqrt{v_{BAx}^{2} + v_{BAy}^{2}} \qquad \textbf{Equation 2.4-12}$$

$$\|V_{AC}\| = \sqrt{v_{ACx}^{2} + v_{ACy}^{2}} \qquad \textbf{Equation 2.4-13}$$

The cross product of the vectors V_{BA} and V_{AC} has a single component in the z direction, which is calculated as:

$$Q = v_{BAx} v_{ACy} - v_{ACx} v_{BAy} \qquad \textbf{Equation 2.4-14}$$

Normalizing the cross product by dividing by the magnitudes of the vectors V_{BA} and V_{AC} gives the sine of the angle between the vectors which is the local change in aircraft course bearing q .

$$\sin q = \frac{Q}{\|V_{BA}\|\|V_{AC}\|} \qquad \textbf{Equation 2.4-15}$$

and

$$q = \sin^{-1}\left(\frac{Q}{\|V_{BA}\|\|V_{AC}\|}\right) \qquad \textbf{Equation 2.4-16}$$

This calculation of q is valid for angles of up to 90 degrees, left or right. For angles from 90 degrees to 180 degrees, left or right, the value of the angle is incorrect, but the sign of the angle is correct. For turn angles greater than 180 degrees, the angle and the sign are incorrect.

The absolute value of $\sin q$ is compared to a threshold to determine whether or not the aircraft is turning. If the aircraft is turning, a positive value of $\sin q$ says the aircraft is turning to the left, a negative value says the aircraft is turning to the right.

In this study, a turn is determined by a nine degree angle (or greater) generated by the two segments drawn from the previous position to the current position and the current position to the next position report. The threshold was determined from observation of several flights in both Indianapolis and Fort Worth ARTCCs. In the future, data smoothing techniques may be employed to further enhance the algorithm changing this threshold angle.

2.4.6.2 Vertical Phase of Flight

The vertical phase of flight for an aircraft was defined as one of three states: level, ascending, or descending. This state was determined by selecting a track data point a parametric number of points (one point in this study) earlier than the point being examined and a track data point a parametric number of points (one point in this study) later than the point being examined. The altitude difference between the earlier point and the later point divided by the time difference between the two points is an estimate of the aircraft's rate of climb or descent. If the absolute value of the measured rate of climb is less than a parametric threshold value (150 feet in this study) the aircraft is considered to be in level flight. If the measured rate of climb is greater than a positive parametric threshold (150 feet in this study) the aircraft is considered to be ascending. If the measured rate of climb is less than a negative parametric threshold (-150 feet in this study), then the aircraft is considered to be descending.

2.5 Trajectory Data Processing and Trajectory Report Generation

Figure 2.5-1 provides a data flow diagram logically describing the data files and processes used to sample the trajectory data and to generate the trajectory reports. This processing consists of the Trajectory Sampling Program (TJS) and the Trajectory Report Generation Program (TRG), discussed in subsections 2.5.1 and 2.5.2.

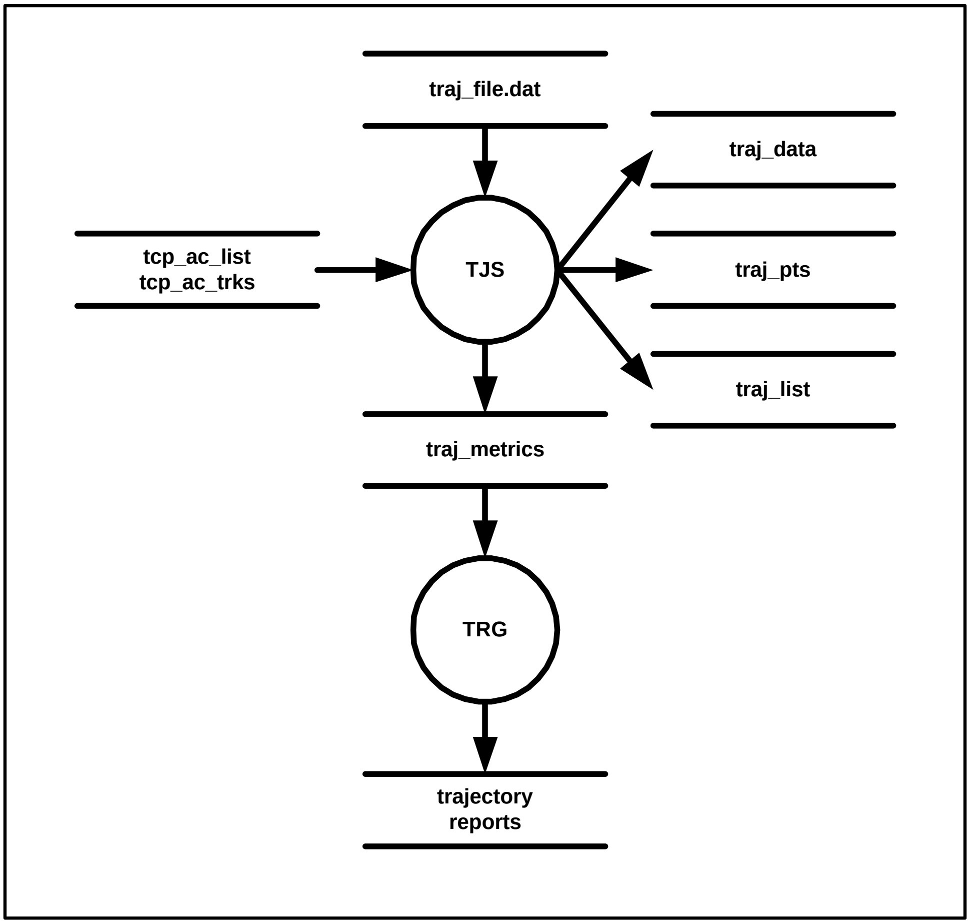

Figure 2.5-1: Trajectory Data Processing and Trajectory Report Generation

2.5.1 Trajectory Sampling Program (TJS)

The Trajectory Sampling Program (TJS) is a C++ program that uses the Oracle Pro*C/C++ Precompiler (Release 8.0) to interface with the Oracle database in the TFM laboratory.

2.5.1.1 Trajectory Sampling

The URET and CTAS trajectory modelers evaluated by this study both compute time-based four-dimensional trajectories. However, they have different design philosophies regarding when these trajectories are calculated.

URET calculates an initial trajectory for each aircraft, then constructs a new trajectory for a given aircraft whenever:

1. A new flight plan or flight plan amendment message is received from the HCS, or new or updated interfacility flight plan information is received from a neighboring URET system.

2. A hold message is received from the HCS that indicates the aircraft is entering or leaving a holding pattern.

3. URET determines that a new trajectory is necessary to reconform an aircraft's trajectory with the aircraft's actual position. This can happen when the HCS track data is found to be a parametric distance (nominally 1.5 to 2.5 nautical miles) from the trajectory or if the current trajectory is older than a parametric value (e.g. 20 minutes).

CTAS, on the other hand, calculates a new trajectory for each aircraft upon receipt of HCS track data each processing cycle.

ACT-250 devised a trajectory sampling technique that is independent of the design approach of either trajectory modeler. The line in Figure 2.5-2 labeled “Track” represents the time line for an aircraft track. The time point labeled T_S represents the initial interpolated track point. The sampling time to start computing metrics for this track is represented by T_0, where

$T_0 = T_S + \text{TRAJ\_DELTA\_TIME}$ **Equation 2.5-1**

TRAJ_DELTA_TIME is a parametric value (40 seconds) which establishes the starting time at a point where the track is more stable.

The trajectories for this example aircraft are presented in Figure 2.5-2 by the time lines labeled $Traj_0$, $Traj_1$, $Traj_2$, and $Traj_3$. The trajectory to be sampled for a particular track sampling time is the trajectory with the latest trajectory build time not exceeding the track sampling time. Selected trajectories were interpolated using techniques similar to the techniques for interpolating tracks described in Section 2.4.4. In Figure 2.5-2, $Traj_0$ would be sampled for sampling time T_0. This point is labeled $T_{0,0}$ and represents the look ahead time of zero seconds for the trajectory sampling time T_0.

Metrics would be computed at the time point labeled T_0 and at the incremented time points $T_{0,1}$ and $T_{0,2}$ where

$T_{i,j+1} = T_{i,j} + \text{TRAJ\_LOOKAHEAD\_TIME}$ **Equation 2.5-2**

TRAJ_LOOKAHEAD_TIME is the parametric sampling interval (300 seconds) for a specific sampling time.

The trajectory sampling process continues until either: the end of the track is reached, the end of the trajectory is reached, or the time exceeds T_0 + TRAJ_LOOKAHEAD_WIN, a parametric input (1800 seconds). Then the next track sampling time T_i will be computed as:

$T_{i+1} = T_i + \text{TRAJ\_SAMPLE\_TIME}$ **Equation 2.5-3**

TRAJ_SAMPLE_TIME is the parametric sampling interval (120 seconds) for sampling a specific track.

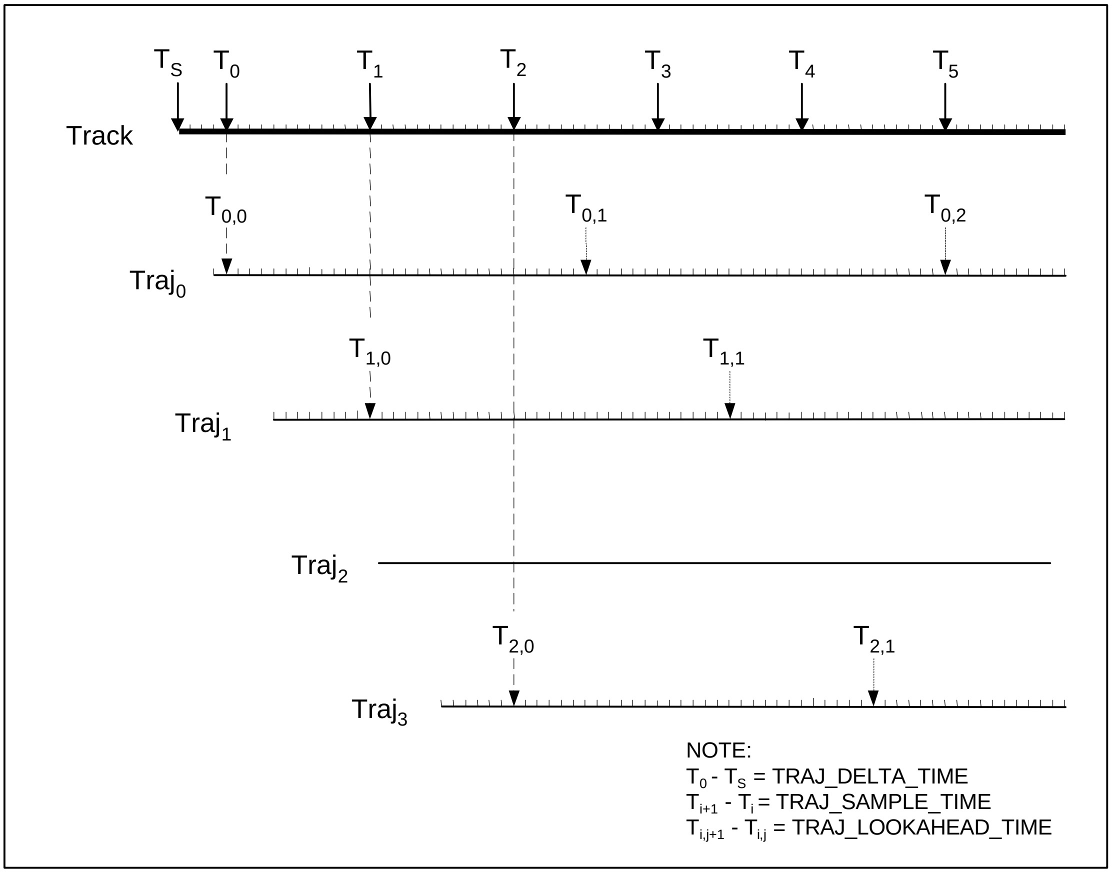

Figure 2.5-2: Interval Based Sampling

2.5.1.2 Estimation of the Metrics

Estimations of the error metrics (the horizontal, longitudinal, lateral, and vertical errors defined in Section 2.2.2) were calculated at a particular time point T as follows. Point A in Figure 2.5-3 represents the actual position of the aircraft at time T, point B represents the predicted position of the aircraft at time T along the trajectory and point C represents the next predicted position along the interpolated trajectory. Line segment AB represents the horizontal error. Point D is defined as the point along the line segment BC at which the angle formed by the line segments BD and DA is a right angle. Then the longitudinal error is represented by the directed line segment BD, and the lateral error is represented by the directed line segment DA.

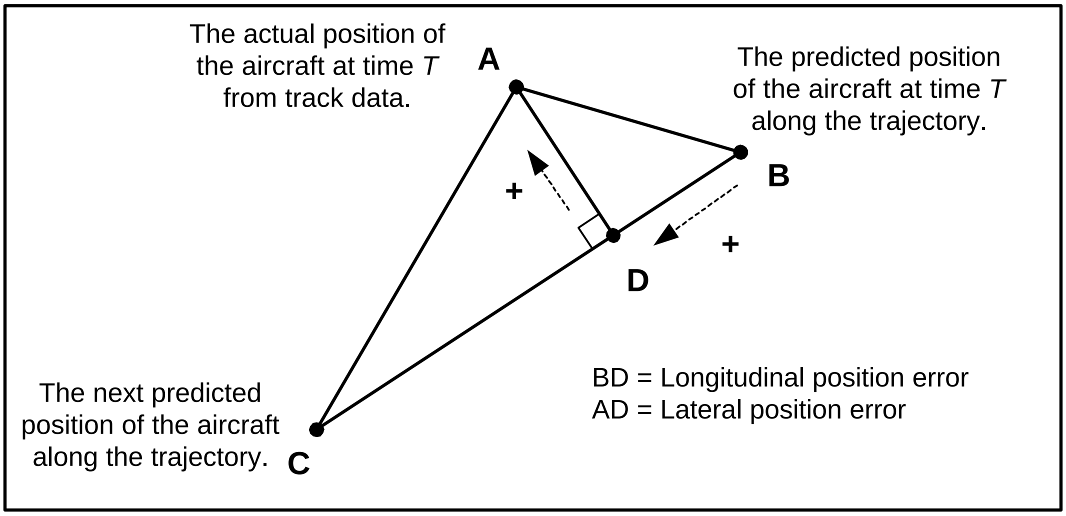

Figure 2.5-3: XY Error Geometry

The method used to calculate these errors is similar to the method used by the URET function GM_PTLINE (see Cale et. al., February 1997, section 3.4.16), described as follows:

As well as the normal case depicted in Figure 2.5-3, there are three special cases: (1) the line BC is parallel to the x-axis, (2) the line BC is parallel to the y-axis, and (3) the points B and C are identical.

If the coordinates of the point A are denoted as (x_A, y_A, z_A), the coordinates of the point B as (x_B, y_B, z_B) and the coordinates of the point C as (x_C, y_C, z_C), then:

1) <u>Normal Case</u>: The slope m of the line BC is then

$$m = \frac{(y_C - y_B)}{(x_C - x_B)} \qquad \textbf{Equation 2.5-4}$$

The slope m' of the line through A perpendicular to BC is the negative reciprocal of m, that is

$$m' = -\frac{1}{m} \qquad \textbf{Equation 2.5-5}$$

The equation of the line through the point A with the slope m' is

$$y = y_A + m'(x - x_A) \qquad \textbf{Equation 2.5-6}$$

The equation of the line through the point B with the slope m is

$$y = y_B + m(x - x_B) \qquad \textbf{Equation 2.5-7}$$

The point of intersection D, denoted as (x_D, y_D), is the simultaneous solution of these two equations:

$$x_D = \frac{y_B - y_A + m^{/} x_A - m x_B}{m^{/} - m}$$ **Equation 2.5-8**

$$y_D = \frac{m m^{/} (x_B - x_A) + m y_A - m^{/} y_B}{m - m^{/}}$$ **Equation 2.5-9**

2) Special Case 1: BC is parallel to the x axis: This is true if and only if $y_B = y_C$. Then the equations for the coordinates of the point D are

$$x_D = x_A$$ **Equation 2.5-10**

and

$$y_D = y_B$$ **Equation 2.5-11**

3) Special Case 2: BC is parallel to the y axis: This is true if and only if $x_B = x_C$. Then the equations for the coordinates of the point D are

$$x_D = x_B$$ **Equation 2.5-12**

and

$$y_D = y_A$$ **Equation 2.5-13**

4) Special Case 3: Points B and C are identical: There is no solution. This case will not occur when the input data for this calculation is valid.

After the coordinates of D have been computed, the longitudinal and lateral errors can be calculated as follows:

The longitudinal error E_{long} is the length of the line BD, which is

$$E_{long} = \sqrt{(x_D - x_B)^2 + (y_D - y_B)^2}$$ **Equation 2.5-14**

The lateral error E_{lat} is the length of the line AD, which is

$$E_{lat} = \sqrt{(x_D - x_A)^2 + (y_D - y_A)^2}$$ **Equation 2.5-15**

The following process was used to determine the signs for the longitudinal and lateral errors. Referring again to Figure 2.5-3 the components of the vectors V_{BA} and V_{BC} are:

$$V_{BA} = \begin{bmatrix} x_A - x_B \\ y_A - y_B \\ z_A - z_B \end{bmatrix} = \begin{bmatrix} n_{BA_x} \\ n_{BA_y} \\ 0 \end{bmatrix}$$ **Equation 2.5-16**

$$V_{BC} = \begin{bmatrix} x_C - x_B \\ y_C - y_B \\ z_C - z_B \end{bmatrix} = \begin{bmatrix} n_{BC_x} \\ n_{BC_y} \\ 0 \end{bmatrix}$$ **Equation 2.5-17**

The scalar dot product of the vectors V_{BA} and V_{BC} is a scalar quantity, which can be calculated:

$$n_{BA_x} n_{BC_x} + n_{BA_y} n_{BC_y}$$ **Equation 2.5-18**

The sign of the longitudinal error was considered positive if this scalar quantity was positive (i.e. track position ahead of trajectory predicted position).

The vector cross product of the vectors V_{BA} and V_{BC} has a single component in the z direction, which can be calculated:

$$n_{BA_x} n_{BC_y} - n_{BA_y} n_{BC_x}$$ **Equation 2.5-19**

The sign of the lateral error was considered positive if the value of this component was positive (i.e. track position to the right of trajectory predicted position).

The vertical error E_{vert} is the signed difference between the altitudes (i.e., the z coordinates) of the two corresponding points from the interpolated track data and the interpolated trajectory data.

$$E_{vert} = z_A - z_B$$ **Equation 2.5-20**

The vertical error is positive when the track position is above the trajectory predicted position.

2.5.2 Trajectory Report Generation

The Trajectory Report Generation (TRG) process is a UNIX shell script and a series of SQL/PL programs that generate several categories of reports, including:

1. Summary and overall statistics on all data including the track and trajectory data.
2. Statistics on the trajectory metrics. There are seven reports for look ahead times equal to zero, 300, 600, 900, 1200, 1500, and 1800 seconds, used in Sections 3.3.1 and 4.3.1.
3. Summary and overall descriptive statistics on the trajectory metrics data, excluding trajectories for which the EARLY_TRAJ_FLAG was set. The EARLY_TRAJ_FLAG flags a trajectory with a build time earlier than the first HCS track report.
4. Descriptive statistics on the trajectory metrics for the seven look ahead times, excluding trajectories for which the EARLY_TRAJ_FLAG was set.

5. A listing of ACID_CID, sample time, trajectory build time, lateral error, longitudinal error, horizontal error, vertical error, and track quality[1] for each look ahead time. This data was used for inferential statistical analysis.
6. Descriptive statistics for the trajectory metrics for each of the seven look ahead times for the horizontal phase of flight including straight and turning, used in Sections 3.3.3 and 4.3.3.
7. Descriptive statistics for the trajectory metrics for each of the seven look ahead times for the vertical phase of flight including level, ascending, and descending, used in Sections 3.3.4 and 4.3.4.
8. Descriptive statistics for the trajectory metrics for each of the seven look ahead times for the following four flight type cases:
   - Overflights
   - Departures
   - Arrivals
   - Internals

   This TRG report was used in Sections 3.3.2 and 4.3.2.
9. Descriptive statistics for the trajectory metrics for each of the given look ahead times for the top ten occurring aircraft types listed in Sections 3.1.5 and 4.1.5 for URET and CTAS, respectively. The use of this TRG report will be left for future studies.
10. Descriptive statistics for the trajectory metrics for each of the given look ahead times for general aviation airlines versus commercial airlines. The use of this TRG report will be left for future studies.

Note: All reports repeated with samples only above 18,000 feet.

2.6 Analysis Methodology

A statistical analysis of the trajectory accuracy of URET and CTAS was conducted. The results of these analyses are presented in Section 3.3 for URET and Section 4.3 for CTAS. The analyses consist of aggregate performance information, such as the number of samples and trajectories analyzed; context related statistics, such as the percentage of flights modeled; and actual trajectory accuracy statistics. For the trajectory accuracy statistics, the analysis is presented in tables delineating the results of inferential statistical tests performed and plots of the mean errors partitioned by selected factors, including look ahead time, phase of flight, and flight type. In addition, complete descriptive statistics for both analyses are contained in Appendices A and B. The following subsections provided additional information on each type of analysis that was conducted.

2.6.1 Aggregate Trajectory Performance Analysis

For the aggregate performance information, counts are reported for the total number of trajectories built, the number of trajectories sampled, and the number of flights processed. The duration of the trajectories and duration of each trajectory analyzed also provide the reader with the magnitude of the analysis coverage. Other aggregate performance information includes the total number of sample points used in the study.

2.6.2 Context Related Trajectory Performance Analysis

The context related statistics provide the reader with knowledge about the scope of the results, including the percentage of valid flights sampled, sampled trajectory age, and ratio of prediction coverage.

[1] Track quality is the percentage of track position reports which have been altered by the RDTRACKS processing.

2.6.2.1 Percentage of Valid Flights Sampled

The first, and probably most important, of the context related statistics is the percentage of valid flights sampled. Two conditions or events were required for a flight to be analyzed: it had to have both flight plan information from the HCS and trajectory prediction data from the DST. Referring to Figure 2.6-1, area "a" defines the valid aircraft flights for analysis. To be valid, an aircraft flight must have (1) a HCS flight plan message, (2) a set of HCS track position reports that have been verified by the RDTRACKS program discussed in Section 2.4.3, and (3) trajectory predictions from the DST. For the events under area "a" in Figure 2.6-1, some time overlap exists between the trajectory prediction and the track position reports. The area "c" includes valid aircraft flights with all the required HCS position data but insufficient trajectory prediction data (i.e., either no trajectory at all or not overlapping in time with the track data). The area "b" in Figure 2.6-1 includes the trajectories built without valid aircraft data, defined as lacking at least one of the HCS data defined above (i.e. flight plan, track data, time overlap, and positional verification).

It is important to quantify these events, since the analysis is based only on area "a". A DST's own bias in building trajectories can influence the trajectory accuracy statistics. In other words, the results are based only on situations when the DST chose to build a trajectory and obviously not on situations where it did not for whatever reason. Therefore, it is important to interpret the trajectory results in context of the trajectories it built. Referring to Figure 2.6-1, the ratio of area "a" to the sum of areas "a" and "c" defines the DST's fraction of valid flights with sampled trajectory prediction. It is reported as the percentage of the valid aircraft flights that have sampled trajectory prediction.

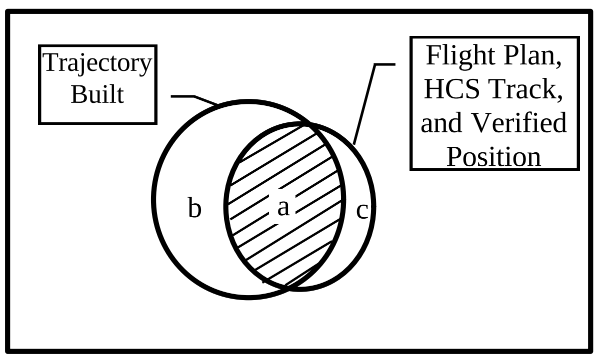

Figure 2.6-1: Trajectory and Aircraft Flight Events Venn Diagram

2.6.2.2 Ratio of Prediction Coverage

Another statistic useful in setting the context of the study estimates the trajectory prediction coverage over the track time analyzed. It is possible for trajectories to exist for a short prediction time with high accuracy while another DST could make predictions for the entire length of HCS track reports with less accuracy. This statistic quantifies this situation. It is defined as the ratio of the total time that the trajectories were predicted and captured for the analysis over the total time that the track was captured for analysis.

Referring to Equation 2.6-1, the trajectory prediction coverage is measured by taking each aircraft in area "a" in Figure 2.6-1 and calculating the difference between its last sampled trajectory's end time and its first sampled trajectory's start time. This difference is then divided by the difference between the end time of its last track report analyzed and the start time of its first track report.

This value will always be less than one, since trajectories are sampled and analyzed starting at 40 seconds past the beginning of the track start time and end with the shorter of the two, either track or trajectory. If a trajectory ends before the track end time, the ratio will be increasingly smaller than one, and if the track ends earlier the ratio will reach a maximum close to one due to the initial 40 seconds delay in sampling.

Equation 2.6—1

$$\text{ratio of prediction coverage} = \frac{(\text{last trajectory's end time - first trajectory's start time})}{(\text{track end time - track start time})}$$

For this analysis, the average and standard deviation of the ratio of prediction coverage is reported, as well as a 95 percent confidence interval around the sample mean. Also a histogram and quantile table (i.e. a table listing the percentiles from 0 to 100) are presented.

2.6.2.3 Sampled Trajectory Age

Another descriptive value that defines the context of the analysis is the age of the trajectory at the look ahead time of zero. Referring to the sampling process defined in Section 2.5.1, the longer a DST retains a trajectory, the older the age of the trajectory at each sampling interval. The age of the trajectory at each sample time is proportional to the frequency trajectories are rebuilt by the DST. In general, a DST that builds trajectories more frequently will have a smaller average trajectory age. Although there may be a correlation between trajectory age and trajectory prediction accuracy, it is also effected by the reasons for the refresh, as well as other factors.

2.6.3 Trajectory Accuracy Analysis

Basic descriptive statistics were calculated for each of the trajectory metrics. These statistics include the average, standard deviation, and maximum and minimum values, for: horizontal error, lateral error, absolute value of lateral error, longitudinal error, absolute value of longitudinal error, vertical error, and absolute value of vertical error. These descriptive statistics are reported for each look ahead time as well as several identified factors. Inferential statistics were used to determine whether the levels of the identified factors were statistically different and had a significant effect on each performance value. For example, at a look ahead time of zero, the hypothesis is tested on whether the mean horizontal error is equivalent in a turn or a straight path. This approach was chosen because of the application of the Central Limit Theorem (CLT), which allows the approximation of a Normal Distribution on a sample mean with a sufficiently large sample size (Devore, 1987). In this study, the sample sizes ranged in the thousands.

For the inferential statistics, three statistical tests were performed[2]:

1. Levene Test which determines if the particular performance value's (e.g. horizontal error) variances are significantly different statistically between the levels (i.e. by look ahead time, different flight types or phases of flight) (Neter, 1996)
2. Welch Test which determines if the particular performance value's sample means are significantly different statistically between the levels (Kelton and Law, 1991)

[2] The three statistical tests defined, Levene, Welch, and Tukey-Kramer, are described in more detail in Appendix A.0. Descriptions of the histograms, box plots, and mean comparison plots (i.e. diamond and circle plots) are also presented in Appendix A.0.

3. Tukey-Kramer Test which determines which of the particular pair or pairs of performance value's sample means are significantly different statistically between the levels (SAS Institute, 1995)

There are many factors which can affect the accuracy of the predictions of the flight path. Section 2.2.3 identifies the factors used in this report; other factors can be analyzed in the future if resources permit.

Table 2.6-1 lists the types of statistical analyses that were performed on each of the identified factors. The analyses included descriptive statistics (tables are presented in Appendix A), or inferential statistics in which hypothesis testing of the means and variances were performed (presented in both Appendix A and summarized in the Sections 3.3 and 4.3 for URET and CTAS, respectively). For several of the factors, both descriptive and inferential statistical analysis was performed. Table 2.6-1 also identifies whether graphical information is presented. Inferential statistics and graphical plots (i.e. histograms and quantile tables) were calculated for a subset of the available look ahead times, including zero, 600, 1200, and 1800 seconds (presented in Appendix A). Also, the Sample Mean Plots are presented in Sections 3.3 and 4.3 for URET and CTAS, respectively, and Sample Standard Deviation Plots are presented in Appendix B. The signed values of the error metrics (e.g. average lateral error) were used for these more exhaustive inferential techniques, since the sample mean acts as a measure of the bias of the trajectory predictions and the standard deviation as a measure of the uncertainty. The absolute value statistics (e.g. average absolute value of lateral error), which are also a useful measure of the uncertainty, have been included in the descriptive statistics reported in Appendix A.

Since the DSTs examined were designed to model IFR aircraft in en route airspace, this study needed a method to generically separate aircraft tracked by the HCS that may have been handed off and were entering a terminal airspace, from other strictly en route flights. The approximate method chosen was to perform two studies, one for all aircraft tracks captured by the HCS and a second performed on HCS track reports above 18,000 feet, which is well above all terminal airspace in the Center's under study. Therefore, all factors including look ahead time were analyzed twice: once with all the sampled track points and then with only sampled track reports above 18,000 feet.

Table 2.6-1: Analysis Summary

Factor For Samples at All Altitudes / Above FL180	Descriptive Statistics	Inferential Statistics	Sample Mean / Std. Dev. Plots	Histograms / Quantiles
Look Ahead Time	Yes	Yes	Yes	Yes
Flight Type	Yes	Yes	Yes	No
Phase of Flight Horizontal	Yes	Yes	Yes	No
Phase of Flight Vertical	Yes	Yes	Yes	No

3. URET Study Results and Observations

The results and observations presented in this section are based on the analysis of over seven hours of data recorded at the Indianapolis ARTCC (ZID). Specific information describing the scenario is presented in Section 3.1. Section 3.2 provides detailed information about one aircraft flight in the study which demonstrates the study's methodology, and Section 3.3 presents the results of the application of the trajectory accuracy metrics to URET.

3.1 Scenario Description

Figure 3.1-1 provides a data flow diagram logically describing the data files and processes used to obtain the flight plan, track, and trajectory data used for the URET analysis. For this study, data was collected from the URET installation at ZID. The source of the data was a Monitor Test and Recording (MTR) file, created at the output of the General Purpose Output Interface Module (GIM), containing the HCS flight plans, flight plan amendments, and track messages sent to URET over a 7.5 hour period on February 27, 1998. The weather data for the same time period was also recorded.

The scenario file, identified as *sn022798.dat* in Figure 3.1-1, was created using the MITRE/CAASD Reverse Host Converge/Merge Process (RHCMP) program (Byrdsong et. al., 1997). The *sn022798.dat* file is an ASCII file containing event records, which are primarily the NAS Host computer messages. These event records contain the time of the event, the event type, the aircraft identifier, and the aircraft's computer identifier followed by the event subfield. The format of these records is defined in Lindsay, 1998. This *sn022798.dat* file was then used as input to both the Flight Plan and Track Data Processing described in Section 2.4, and to URET D3A (specifically, URET Release D3A_R3_P2) in the WJHTC TFM laboratory.

The trajectory information was recorded by URET's Data Recorder program in binary format. The trajectory data is first parsed into a large ASCII file by MITRE's Data Collection Post Processor, DCPP, (Byrdsong et. al., 1997). This file, *ssg_file*, still needs to be parsed further and converted to a generic format. The *ssg_file* is input into a program composed of a UNIX shell script and C++ program called *up_scr*. This program parses the trajectory data into a generic ASCII file called *traj_file.dat*, which was input to the Trajectory Data Processing described in Section 2.5. The formats of the *ssg_file* and the *traj_file.dat* files are described in WJHTC/ACT-250, 1999.

Tables 3.1-1 and 3.1-2 summarize the characteristics of the airspace and the aircraft flights through the airspace, respectively, for the subject scenario.

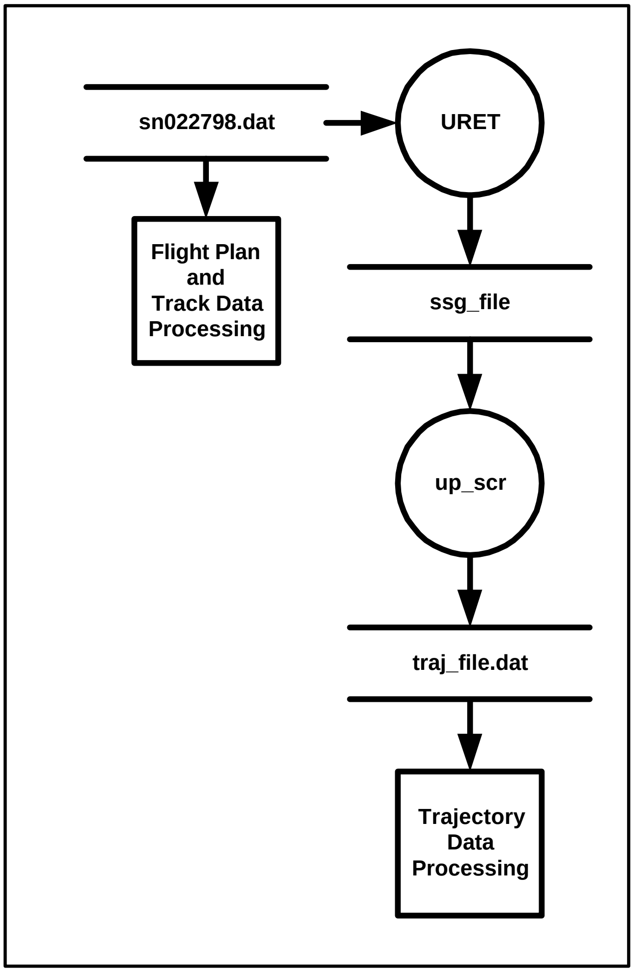

Figure 3.1-1: URET Data Sources

3.1.1 Airspace Definition

Table 3.1-1 summarizes the spatial and time boundaries of the ZID data sample used.

Table 3.1-1: ZID Airspace Definition for URET Study

Airspace	Indianapolis Center (ZID)
Altitude	0 to 60,000 feet
Horizontal boundaries	Defined by the high altitude sectors
Date	February 27, 1998
Start time	12:01:31 UTC (6:01 a.m. local time)
End time	19:33:10 UTC (1:33 p.m. local time)
Duration	07:31:39 or 27,099 seconds

3.1.2 Aircraft Counts

Table 3.1-2 delineates the counts of aircraft flights in the sample of air traffic analyzed.

Table 3.1-2: Aircraft Counts for URET Study

Total number in sample (IFR)	2656
Number excluded	150 (5.65 %)
Number processed	2506 (94.4 % of total)
Number of airliners	1913
Number of General Aviation aircraft	593
Number of jets in the top 20 aircraft	15
Number of turboprops in the top 20 aircraft	5
Number of piston aircraft	0
Average length of track	34.7 minutes, 2082 seconds, or 174 position reports
Number of overflights	1115 (44.5 %)
Number of departures	692 (27.6 %)
Number of arrivals	630 (25.1 %)
Number of internal flights	69 (2.8 %)

3.1.3 Excluded Flights

In measuring the accuracy of track predictions, the true positions of the aircraft are assumed to be the positions reported by the HCS. For some aircraft, it is clear that the HCS reported positions are not correct. Track processing algorithms were used to correct the position data where possible, as described in Section 2.4. When it was not possible to correct the data, the individual tracks and in some cases entire flights were deleted from the scenario being examined, as discussed in the following sections. Statistics were collected on an aircraft flight only if both a track and a set of predicted trajectories were available. For this analysis of URET, there were three categories of excluded aircraft totaling 150 flights that were deleted from the original set of 2656 IFR flights (a reduction of 5.65 %).

3.1.3.1 Military Flights

Since it is often not possible from flight plan data to accurately predict the flight paths of military flights, which usually are doing either gunnery practice or aerial re-fueling maneuvers, military flights were excluded from the analysis. This was done by selecting out all of the flights which had a call sign containing more than three leading alphabetic characters (e.g., ANVIL, CODER, RACER, SABER, STEEL). Although this is not an exact definition of military aircraft, it was considered to be sufficient for this study. 79 military flights were excluded.

3.1.3.2 Non-initialized Flights

As discussed in Section 2.4, sometimes the HCS processing algorithms are unable to establish a consistent track for the aircraft. There were 18 flights excluded for this reason.

3.1.3.3 Uncertain Position Flights

The processing of the HCS track data requires correcting some of the track reports which are clearly in error. For example, as discussed in Section 2.4.3, sometimes the same XY coordinates are repeated even though the aircraft has moved between the radar reports. Now in some cases the corrected track reports are substantially different from the original aircraft positions reported by the HCS. This situation implies that we, the experimenters, do not know the true position of the aircraft. Flights having a corrected track position report substantially different from the original position report were deleted (53 of these flights were excluded).

3.1.4 Truncated Flights

Often in the HCS track reports, several tracks reports are missing or have bad data. The position of the aircraft during the gap is unknown. If the gap is short, the missing track reports can be interpolated. When a large gap in the track data occurs, the track positions after the gap are discarded. Of the 452,976 radar track position reports, 15,756 or 3.6 % were discarded by truncating the tracks after missing or bad data.

Measurements of trajectory prediction errors were made on aircraft either already in the ZID airspace or approaching the ZID airspace and about to be in the airspace. Measurements were not made on aircraft after they left ZID airspace. That is, no measurements were made on the portions of the tracks outside ZID when the aircraft were flying away from the ZID airspace. 17.2% of the interpolated track reports were not used for this reason.

3.1.5 Aircraft Mix

The majority of the aircraft analyzed in this study are commercial airliners. The top 10 aircraft type account for 1358 of the 2506 flights, or 54.2 % of the total; the top 20 aircraft account for 1746 of the 2506 flights, or 69.7 % of the total. A histogram depicting the frequency of occurrence of the top 20 aircraft is provided in Figure 3.1-2. The aircraft are identified by their FAA type designators. Of the top 20 aircraft, 15 are jets and five are turboprops. Table 3.1-3 lists the aircraft manufacturers and model names of the top 10 aircraft. All of the top 10 aircraft are jets except for the EMB 120.

Table 3.1-3: URET Scenario Aircraft

RANK	FAA TYPE IDENTIFIER	MANUFACTURER / MODEL	NUMBER OF FLIGHTS	PERCENTAGE OF FLIGHTS
1	DC9	McDonnell-Douglas DC9	224	8.94 %
2	B727	Boeing 727	186	7.42 %
3	B73B	Boeing 737-300/400/500	182	7.26 %
4	CARJ	Canadair Bombardier Regional Jet	152	6.07 %
5	B757	Boeing 757	143	5.71 %
6	MD80	McDonnell- Douglas MD80	131	5.23 %
7	MD88	McDonnell-Douglas MD88	122	4.87 %
8	B73A	Boeing 737-200	87	3.47 %
9	E120	Embraer EMB 120	78	3.11 %
10	B737	Boeing 737-200	53	2.11 %

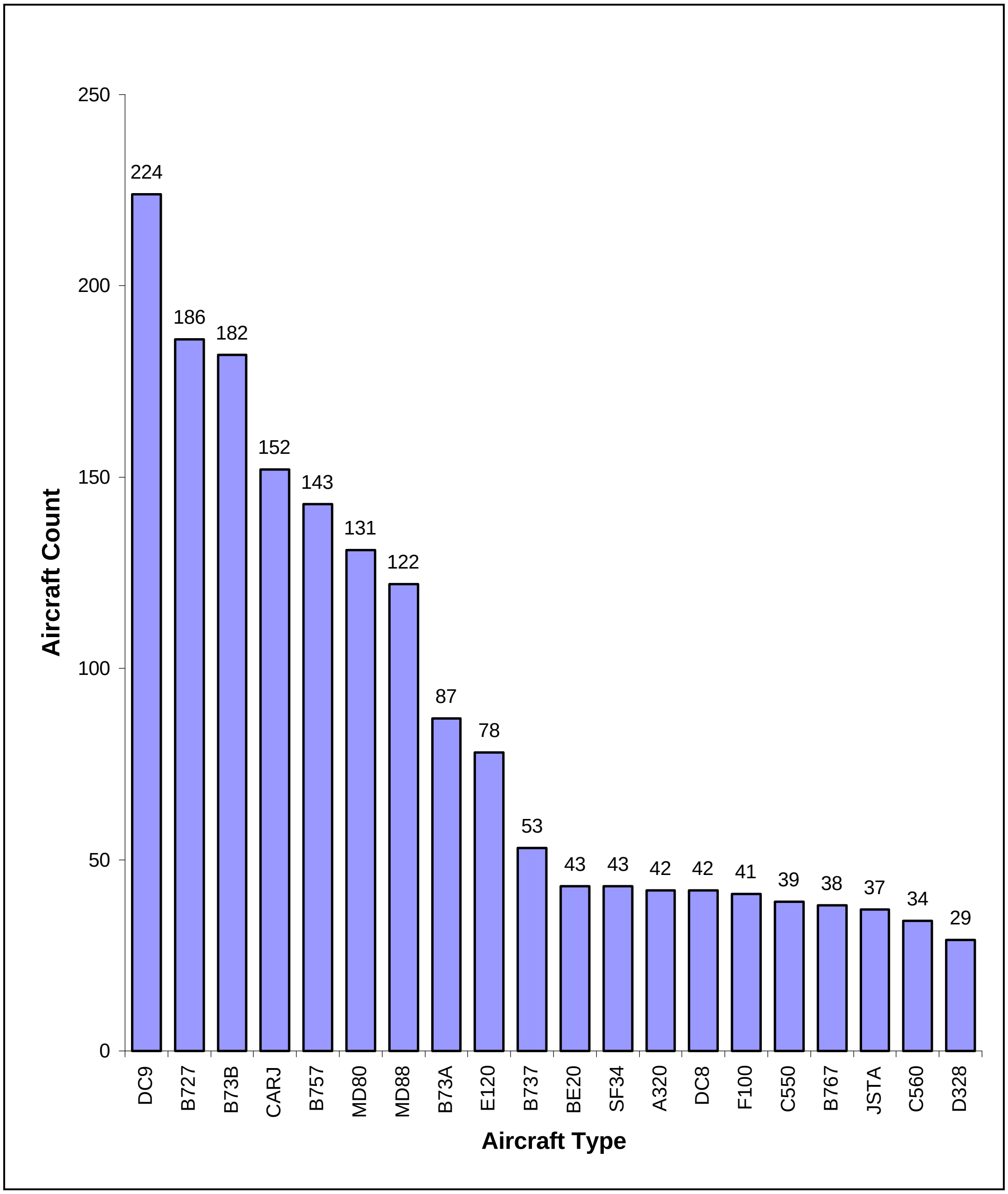

Figure 3.1-2: Top 20 Aircraft Frequency Histogram - ZID Data

3.2 Observations

This section presents observations made during analysis of the data, which provide detailed information about a specific aircraft flight in the URET study. These observations are included before the results so that the reader can better understand the methodology, and therefore better understand the statistics and data presented in Section 3.3. While each observation details a typical flight, the errors are not necessarily representative of common occurrences. Appendix C provides additional anomalous flights, which were selected to verify the methodology and to examine trajectory accuracy errors with URET.

3.2.1 URET1

In this example, a Boeing 737 commercial airliner departed Baltimore-Washington International (BWI) enroute for Chicago's Midway Airport (MDW). The filed route was J149 and the filed altitude was flight level (FL) 350. This route was an overflight through the northeast part of the ZID airspace. The filed route from BWI to MDW is shown in Figure 3.2-1 with selected waypoints illustrated as small circles.

3.2.1.1 Track Data

The HCS acquired the radar track while the aircraft was in West Virginia (Washington Center, ZDC) on J149 heading west towards ZID. The HCS tracked the aircraft until it left ZID and entered the Chicago Center (ZAU) airspace heading towards Fort Wayne (FWA) on J149. The track data extends all the way to the Goshen VORTAC (GSH); however, no trajectory accuracy measurements were made after the aircraft left the ZID airspace. The track is shown in Figure 3.2-1. The track and the Flight Plan route are coincident.

The aircraft followed its filed route and filed altitude until a flight amendment was submitted to descend the aircraft from FL 350 to FL 310. After the amendment was submitted, there was an altitude hold at FL 350 for about a minute. Then the aircraft was cleared to the interim altitude of FL 330. The aircraft paused briefly at FL 330, and then, after being cleared, continued down to FL 310. The aircraft exited the ZID airspace at FL 310. Its Top of Descent (TOD) from FL 310 was outside of ZID. The altitude profile is shown in Figure 3.2-2.

The radar position reports supplied by HCS were reasonably consistent. Of the 244 position reports, 10 were defective and had to be fixed. The first track report had zero altitude and was discarded. There were five stationary position reports, which repeated the previous position report. The XYZ coordinates for these reports were replaced by interpolated values. There were four position reports which had zero altitude and one position report which was both stationary and had zero altitude. These reports were replaced by interpolated values as well.

3.2.1.2 Trajectory Data

The track time and the time lines for the eight trajectories recovered for this aircraft are presented in Figure 3.2-3. The time line for the track is labeled "Track." The time lines for the trajectories are labeled with the trajectory's build time. The first three of these trajectories (the 45728, 45729, and 47218 trajectories) were built before the first track point at time 47230. The sample points for calculating the trajectory accuracy metrics are shown by arrows drawn from the track time line to the latest trajectory available at that sample time. The first sample time was 47270 (40 seconds after the first track point). This sample used the 47230 trajectory which was built with the first track point. Of these eight trajectories three were sampled: the 47230, 49062, and 49194 trajectories.

The three trajectories have been plotted in Figures 3.2-1 and 3.2-2. In the plan view (Figure 3.2-1), it can be seen that the trajectories are coincident with the filed route when the aircraft is approaching and within the ZID airspace. In the altitude profile plot (Figure 3.2-2), it can be seen that the trajectories differ from the track data near the TOD.

The trajectories plotted all start with a data point, which is sampled for the error measurements. Previous trajectory points have been discarded because they are not needed for the metric calculations. Up to two minutes of initial trajectory data may be discarded. For example, the first data point plotted for Trajectory 3 is at 49,310 seconds, although the trajectory was built at 49,194 seconds.

3.2.1.3 Metrics

Table 3.2-1 presents the trajectory metrics calculated for this aircraft. The longitudinal and lateral errors are in nautical miles; the vertical errors are in feet. As discussed in Section 2.5.1, a sample is taken 40 seconds after the start of track and then repeated each two minutes until either the track ends, the trajectory ends, or the track leaves the center. At each sample time, the distance between the track and trajectory was calculated at the current time and at look ahead times of 5, 10, 15, 20, and 30 minutes into the future. That is, measurements were made at look ahead times of 0, 300, 600, 900, 1200, 1500, and 1800 seconds. The metrics were not computed after time 49430 because the aircraft departed the ZID airspace at 49,550 seconds. The data in the table shows that both the longitudinal and lateral errors were small even at the higher look ahead times. The plot of the track and trajectory data in Figure 3.2-1 shows that the lateral errors are negligible. (The plot does not show the longitudinal errors.)

The vertical profile plot in Figure 3.2-2 shows that near the TOD there are differences in altitude between the predicted trajectories and the actual track flown. The first trajectory predicts an initial TOD at a time of 49,350 seconds and an initial Bottom of Descent (BOD) at an altitude of 31,000 feet and a time of 49,500 seconds. The actual (track) initial TOD was at 49,080 and the actual (track) initial BOD was at 49,370. The predicted TOD was updated to 49,100 by the second predicted trajectory when a Flight Plan Amendment was received. The second trajectory descended the aircraft to an interim altitude of 33,000 feet, held it there for four minutes, and then descended it to 31,000 feet starting at 49,420 reaching 31,000 feet at 49,500, and then it had a final descent, leaving 31,000 at 49,910. The track did not hold at 33,000 feet. The plot of the third trajectory flies the aircraft at 31,000 feet, coincident with the track, passing out of the ZID airspace before descending.

The inaccurate predictions of the TOD and the interim altitude hold produce errors in the predicted altitudes. Error measurements are made every 60 seconds (for some look ahead time). Measurements made at 49,190, 49,250, 49,310, 49,370, and 49,430 seconds show large altitude errors. All of the large altitude prediction errors except one are based on Trajectory 1. The other large altitude error is based on Trajectory 2. The errors have been listed in Table 3.2-1. The time of measurement is the sum of the sample time and the look ahead time. Figure 3.2-2 shows the differences in altitude between the track data and the predicted trajectories which produce these altitude errors.

The largest error (3629 feet) occurred at 49,370 when the aircraft had leveled off at 31,000 feet and it had been predicted to be just past its initial TOD, descending from 35,000 feet. This measurement was made for a look ahead of 15 minutes.

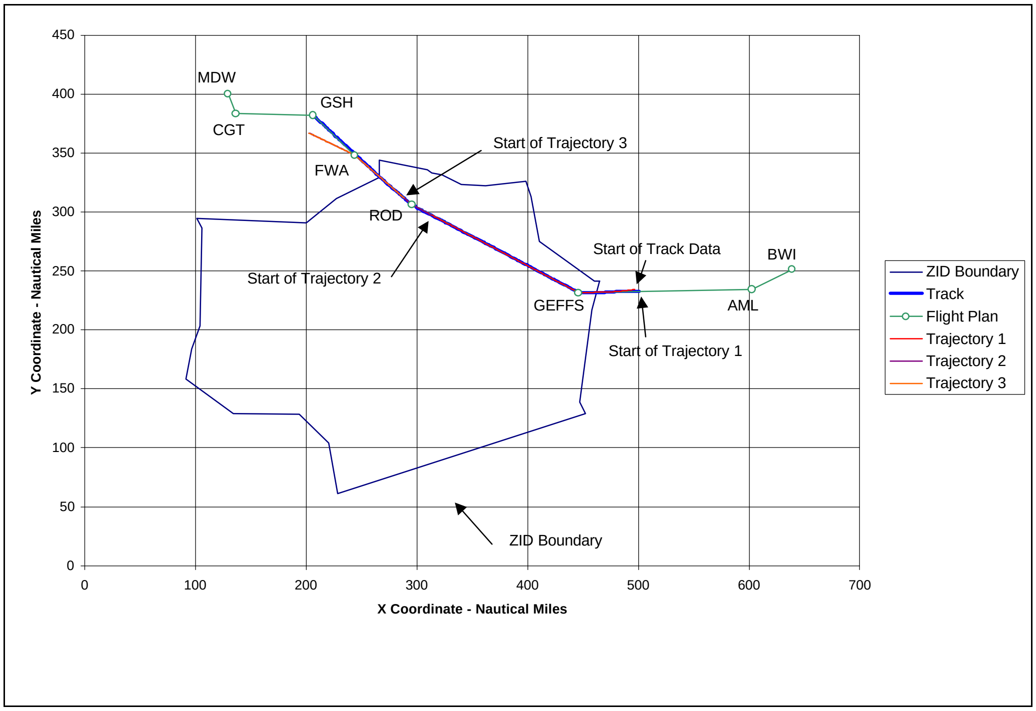

Figure 3.2-1: Aircraft Track and Route

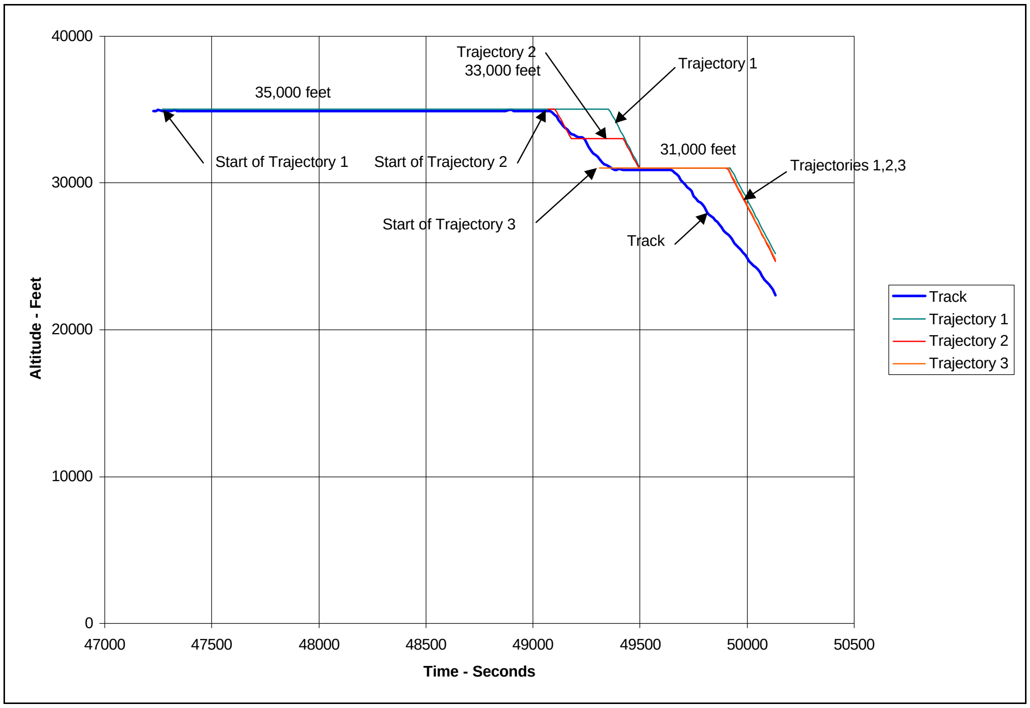

Figure 3.2-2: Altitude Vs. Time

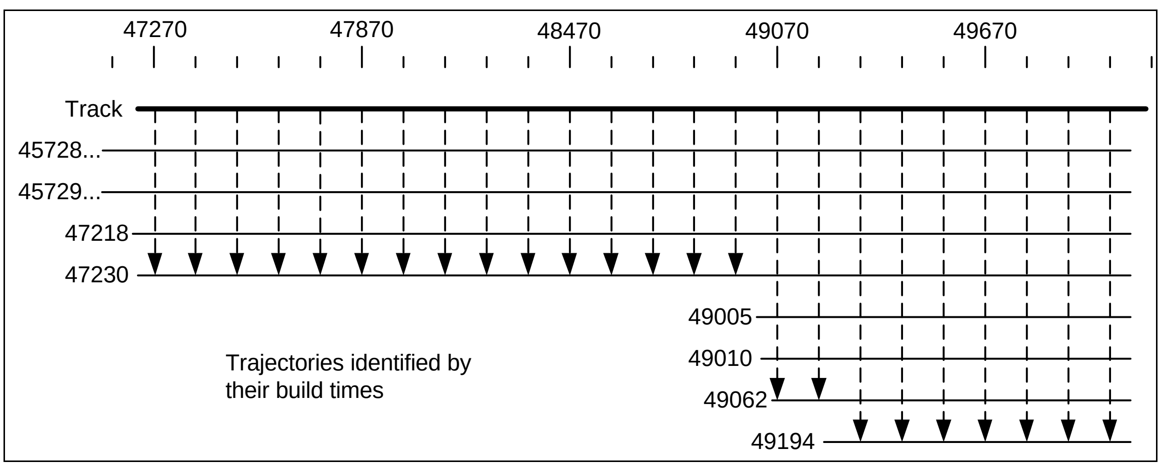

Figure 3.2-3: Sampled Trajectories

Table 3.2-1: Trajectory Metrics (1 of 2)[3]

Sample Time	Traj No	Traj Build Time	Look Ahead Time	Long Error	Lat Error	Vert Error
47270	1	47230	0	-0.23	-1.63	-100.00
			300	-0.50	-0.12	-100.00
			600	-0.71	0.13	-100.00
			900	-0.85	0.18	-100.00
			1200	-0.60	0.21	-100.00
			1500	-1.16	-0.09	-100.00
			1800	-0.52	-0.25	-100.00
47390	1	47230	0	-0.38	-0.42	-100.00
			300	-0.67	-0.05	-100.00
			600	-0.81	0.09	-100.00
			900	-0.92	0.22	-100.00
			1200	-0.52	0.26	-100.00
			1500	-0.49	-0.23	-33.00
			1800	-0.16	-0.35	-1733.00
47510	1	47230	0	-0.46	-0.11	-100.00
			300	-0.62	0.24	-100.00
			600	-0.90	0.14	-100.00
			900	-0.13	0.36	-100.00
			1200	-0.81	-0.03	-100.00
			1500	-0.55	-0.09	-100.00
			1800	0.54	-0.14	-3400.00
47630	1	47230	0	-0.56	-0.09	-100.00
			300	-0.66	0.07	-100.00
			600	-0.91	0.12	-100.00
			900	-0.55	0.30	-100.00
			1200	-1.08	-0.19	-100.00
			1500	-0.39	-0.28	-956.00
			1800	1.03	-0.27	-2061.60
47750	1	47230	0	-0.70	0.12	-100.00
			300	-0.84	0.16	-100.00
			600	-0.85	0.11	-100.00
			900	-0.54	0.13	-100.00
			1200	-0.44	-0.20	-100.00
			1500	-0.41	-0.39	-2300.00
47870	1	47230	0	-0.71	0.13	-100.00
			300	-0.85	0.18	-100.00
			600	-0.60	0.21	-100.00
			900	-1.16	-0.09	-100.00
			1200	-0.52	-0.25	-100.00
			1500	0.74	-0.20	-3629.08
47990	1	47230	0	-0.81	0.09	-100.00
			300	-0.92	0.22	-100.00
			600	-0.52	0.26	-100.00
			900	-0.49	-0.23	-33.00
			1200	-0.16	-0.35	-1733.00
			1500	1.27	-0.10	-400.57

[3] In this chart, longitudinal and lateral error are reported in hundredths of nautical miles, and the vertical error is reported in hundredths of feet. The precision of the input HCS altitude data is reported to the nearest 100 feet, the apparent difference is simply an artifact of the track report processing.

Table 3.2-1: Trajectory Metrics (2 of 2)

Sample Time	Traj No	Traj Build Time	Look Ahead Time	Long Error	Lat Error	Vert Error
48110	1	47230	0	-0.90	0.14	-100.00
			300	-0.13	0.36	-100.00
			600	-0.81	-0.03	-100.00
			900	-0.55	-0.09	-100.00
			1200	0.54	-0.14	-3400.00
48230	1	47230	0	-0.91	0.12	-100.00
			300	-0.55	0.30	-100.00
			600	-1.08	-0.19	-100.00
			900	-0.39	-0.28	-956.00
			1200	1.03	-0.27	-2061.60
48350	1	47230	0	-0.85	0.11	-100.00
			300	-0.54	0.13	-100.00
			600	-0.44	-0.20	-100.00
			900	-0.41	-0.39	-2300.00
48470	1	47230	0	-0.60	0.21	-100.00
			300	-1.16	-0.09	-100.00
			600	-0.52	-0.25	-100.00
			900	0.74	-0.20	-3629.08
48590	1	47230	0	-0.52	0.26	-100.00
			300	-0.49	-0.23	-33.00
			600	-0.16	-0.35	-1733.00
			900	1.27	-0.10	-400.57
48710	1	47230	0	-0.81	-0.03	-100.00
			300	-0.55	-0.09	-100.00
			600	0.54	-0.14	-3400.00
48830	1	47230	0	-1.08	-0.19	-100.00
			300	-0.39	-0.28	-956.00
			600	1.03	-0.27	-2061.60
48950	1	47230	0	-0.44	-0.20	-100.00
			300	-0.41	-0.39	-2300.00
49070	2	49062	0	-0.44	-0.25	-100.00
			300	0.09	-0.20	-2033.00
49190	2	49062	0	-0.33	-0.35	267.00
			300	0.49	-0.10	-238.11
49310	3	49194	0	0.05	-0.14	600.00
49430	3	49194	0	0.51	-0.27	-100.00

3.3 Results

After running URET Delivery 3A with the 7.5 hour scenario file described in Section 3.1, a total of 16,631 trajectories were sampled out of 40,894 trajectories. The sampled trajectories were from 2436 flights. Therefore, each one of these flights on average had 6.8 trajectories analyzed. The average duration of these trajectories is 57 minutes with standard deviation of 39 minutes. The sampling process reduced the trajectory to the portion where both HCS track data and the predicted trajectory overlap in time, so the duration of the trajectory actually analyzed was reduced to approximately 29 minutes on average, with a standard deviation of 18 minutes.

To set the context of the study as defined in Section 2.6.2.1, the counts of the event areas illustrated in Figure 2.6-1 are listed in Table 3.3-1 below. Referring to Figure 2.6-1, the ratio of area “a” to the sum of areas “a” and “c” defines URET’s fraction of valid flights with sampled trajectory prediction. For URET, 97.2 percent of the valid aircraft flights had sampled trajectory prediction.

Table 3.3-1: Valid Track and Trajectory Counts for URET Scenario

	Valid HCS Flight Data	Insufficient Valid HCS Flight Data	Total Flights With Trajectories
Trajectory	2436 (a)	1296 (b)	3732 (a +b)
Insufficient Trajectory	70 (c)		
Total Valid Flights	2506 (a + c)		

As defined in Section 2.6.2.2, another statistic useful in setting the context of the study estimates the trajectory prediction coverage over the track time analyzed. For URET, each analyzed flight had an average of 96.6 percent of prediction coverage with a standard deviation of 6.1 percent. Referring to Figure 3.3-1 and the Quantiles in Table 3.3-2, the distribution decreases very sharply, making a narrow 95 percent confidence interval around the mean between 96.4 to 96.9. The maximum ratio of prediction coverage for URET was 99.4 percent and the minimum was 2.9 percent.

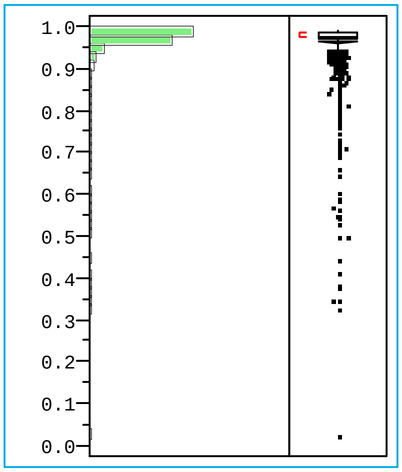

Figure 3.3-1: URET’s Distribution of Ratio of Coverage Statistic

Table 3.3-2: Quantile Table of Ratio of Prediction Coverage

Quantile Label	Percentile	Value
maximum	100.00%	0.99434
	99.50%	0.99246
	97.50%	0.99024
	90.00%	0.98813
quartile	75.00%	0.98491
median	50.00%	0.97938
quartile	25.00%	0.97037
	10.00%	0.94964
	2.50%	0.84657
	0.50%	0.5
minimum	0.00%	0.02913

As described in Section 2.6.2.3, another descriptive value that defines the context of the analysis is the age of the trajectory at the look ahead time of zero. For URET, trajectories are built when the HCS track positions are outside thresholds (referred to as conformance boxes) around the trajectory centerline, when certain messages are received from the HCS, or every 20 minutes. This study's sampled URET trajectories have an average trajectory age of approximately four minutes with a standard deviation of 5.1 minutes.

As discussed above, URET builds trajectories every 20 minutes maximum and often earlier depending on the HCS track positions. The build time in seconds combined with the aircraft identifier string and HCS CID should uniquely represent a particular trajectory. However, there are instances that an aircraft has multiple trajectories with common build times. This is an anomaly of URET's data recording software, which runs in parallel to the URET processes but apparently has a lower priority on machine resources. The anomaly occurs when the data recorder builds up a queue in its processing and gets behind the data being stored in the URET databases. If more than one trajectory is in the queue for a particular flight, the time stamps of the trajectories utilized for the build time can get duplicated creating common trajectory build times. The solution applied was to add one second to the trajectory build time (i.e. sequentially by recording order) in these instances. For the scenario in this study, around 10 percent of the 40,894 URET trajectories needed this adjustment. Once again, the adjustment was only to the build time and was only changed by one second.

The actual trajectory metrics and sampling process is defined in Section 2.5.1. For this 7.5 hour ZID scenario, 138,532 samples were taken against the 16,631 trajectories discussed above. Each sample consisted of spatial prediction error measurements including horizontal error, lateral error, longitudinal error, and vertical error. These measures are reported as a function of different look ahead times from zero to 30 minutes in the future, so the trajectory prediction performance includes the spatial prediction errors partitioned by look ahead time. As a review, look ahead time is the predicted time into the future measured from the sample start time for that particular flight. In this study increments of five minutes were used up to a look ahead time of 30 minutes into the future. In other words, if the flight had both a sampled trajectory and sufficient HCS track reports for the full range of time overlap, error measurements would be calculated at zero, five, 10, 15, 20, 25 and 30 minutes into the future.

Table 3.3-3 lists the types of statistical analyses that were performed on each of the identified factors. The analyses include either descriptive statistics in which simple tables are presented, inferential statistics in which hypothesis testing of the means and variances were performed, or both. This table also lists whether graphical information was presented with references to the appropriate section number. Inferential statistics and graphical plots (i.e. histograms and quantile tables) were calculated for a subset of the available look ahead times, including zero, 600, 1200, and 1800 seconds. The signed values of the error metrics (e.g. average lateral error) were used for these more exhaustive inferential techniques, since the sample mean acts as a measure of the bias of the trajectory predictions and the standard deviation as a measure of the uncertainty. The absolute value statistics (e.g. average absolute value of lateral error), which are also a useful measure of the uncertainty, have been included in the descriptive statistics reported in Appendix A.1.

Table 3.3-3: URET Analysis Summary

Factor For Samples at All Altitudes / Above FL180	Descriptive Statistics	Inferential Statistics	Histograms / Quantiles	Section Number
Look Ahead Time	Yes	Yes	Yes	3.3.1
Flight Type	Yes	Yes	No	3.3.2
Phase of Flight Horizontal	Yes	Yes	No	3.3.3
Phase of Flight Vertical	Yes	Yes	No	3.3.4

3.3.1 Analysis of Look ahead time on Trajectory Accuracy

The main factor analyzed in this study was look ahead time, defined in Section 2.2.3.3. One would expect look ahead time to have a statistically significant effect on performance, but the magnitude of the effect is also of interest. A complete table of the spatial prediction error statistics are presented at the look ahead times of zero, 300, 600, 900, 1200, 1500, and 1800 seconds (i.e. zero to 30 minutes) in Appendix A.1. The focus of the following analysis is on the signed error for lateral, longitudinal, horizontal, and vertical errors at the look ahead times of zero, 600, 1200, and 1800 seconds. This analysis includes an example set and summary results of several tables of statistical information provided by the SAS-JMP Software package (SAS Institute, 1995). They are used to evaluate the error data categorized by look ahead time and in the later sections by horizontal and vertical phase of flight. Complete tables for the URET data are provided in Appendix A.1. The tables present test results for unequal variance including the Levene Test and the Welch Anova Test. They also include a pairwise means comparison, referred to as the Tukey-Kramer Honestly Significant Difference (HSD) Test. Graphical plots present a comparison of means with a quantile box, a plot of the means at look ahead time versus error, and a plot of means using the Tukey-Kramer criteria.

3.3.1.1 Samples at all altitudes

The sample variance of the horizontal error from the four look ahead times are compared first by a Levene Statistical test (Neter, 1996). Referring to Table 3.3-4, this statistical test determines if the hypothesis of equal variances can be rejected. The hypothesis can be rejected in this case, since the variances are significantly different. From Table 3.3-4, the variance of horizontal error is increasing as the look ahead time increases.

Table 3.3-4: Tests for Equal Variances and Tests for Equal Means

Tests that the Variances are Equal (Horizontal Error)[4]				
Level (seconds)	Count	Std Dev (nm)	MeanAbsDif To Mean (nm)	MeanAbsDif To Median (nm)
0	35928	1.08	0.71	0.69
600	23964	5.47	3.66	3.36
1200	13836	8.89	5.82	5.39
1800	6444	10.90	7.01	6.49
Test	F Ratio	Deg of Freedom	DF Den	Prob>F
Levene	7382.12	3	80168	0.0000
Welch Anova testing Means Equal, allowing Std's Not Equal				
	F Ratio	Deg of Freedom	DF Den	Prob>F
	8172.26	3	18809	0.0000

Next, the sample mean for each look ahead time is compared. Referring to Table 3.3-4, the Welch test is applied which compares distributions with different variances and sample sizes. It tests whether all the group means are equal. For the horizontal error at different look ahead times, the Welch Test provides evidence to reject the hypothesis that these mean errors are equal. In Figure 3.3-2, diamonds are drawn around each mean representing the 95 percent confidence interval (in this case, the diamonds are flat and look more like heavy lines due to the large range between the group means). These confidence intervals show an increase in the average horizontal error from zero to 1800 seconds look ahead time of approximately 9.0 nautical miles, from 1.2 nautical miles to 10.2 nautical miles.

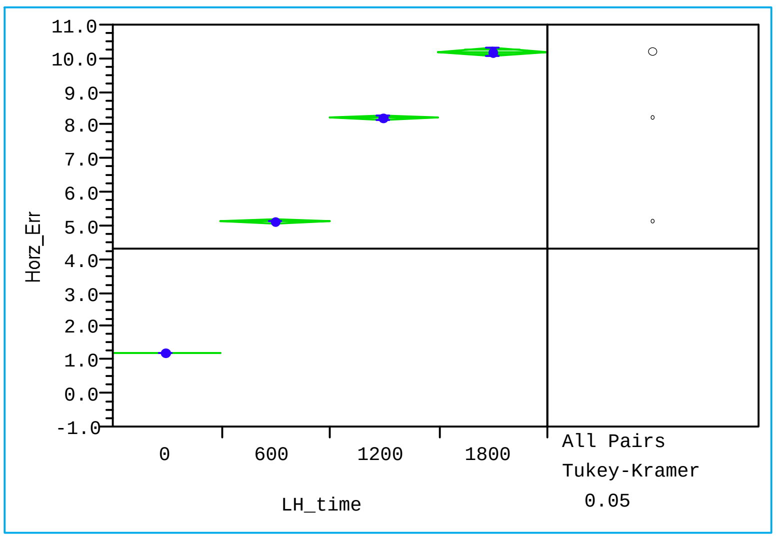

Figure 3.3-2: Sample Mean Comparison of Horizontal Error at Four Look Ahead Times[5]

---

[4] Mean Absolute difference to mean and median are intermediate calculations in the Levene Test described in Appendix A.0.

[5] Normally, the height of the diamond is the length of the confidence interval and the width is proportional to the sample size. In this study, the width has been set equal for all sample sizes.

The lower portion of Table 3.3-5 presents the results of a third statistical test, called the Tukey-Kramer Test, that compares all pairs of means and holds the Type I error at 0.05 for the entire test. It has the exact Type I error if the sample sizes are equal, and is conservative if they are not, which is the case in this study. The horizontal error at the four look ahead times is significantly different between all pairs. The Tukey-Kramer Test provides a distance referred to as the Least Significant Difference (LSD)[6] that can be subtracted from the absolute difference of each pair of means. If the result is positive, the absolute difference of the means is greater than LSD, and the pair of means is significantly different. If the result is negative, the LSD is greater, and the pair is not significantly different. The upper portion of Table 3.3-5 lists the pairwise differences of the sample means for the various look ahead times. All these pairwise comparisons of the means of the horizontal error at the different look ahead times were significant.

The right side of Figure 3.3-2 presents a graphical form of the Tukey-Kramer Test. Too small to be drawn in some cases, it constructs circles around the sample means with a diameter approximately equal to the 95 percent confidence interval. However, this interval is expanded to account for the comparison of all pairs. In short, if the circles overlap the means are not considered significantly different; if they do not overlap, the means are considered significantly different. The circles drawn in Figure 3.3-2 are not overlapping at all, illustrating the numerical results that all the means are different.

Table 3.3-5: Statistical Comparison of All Means (Horizontal Error)

Means Comparisons				
Dif=Mean[i]-Mean[j]	1800	1200	600	0
1800	0.00	1.92	5.06	8.96
1200	-1.92	0.00	3.14	7.04
600	-5.06	-3.14	0.00	3.90
0	-8.96	-7.04	-3.90	0.00
Comparisons for all pairs using Tukey-Kramer HSD				
q* = 2.56909	Alpha= 0.05			
Abs(Dif)-LSD	1800	1200	600	0
1800	-0.26	1.70	4.85	8.76
1200	1.70	-0.18	2.98	6.90
600	4.85	2.98	-0.13	3.78
0	8.76	6.90	3.78	-0.11
Positive values show pairs of means that are significantly different.				

[6] LSD is proportional to the square root of the sum of the squared product of q* and the standard error of both means being compared. The q* value is a quantile similar to the t value of a Student t distribution but expanded to account for the alpha being held for the entire set of comparisons (SAS Institute, 1995).

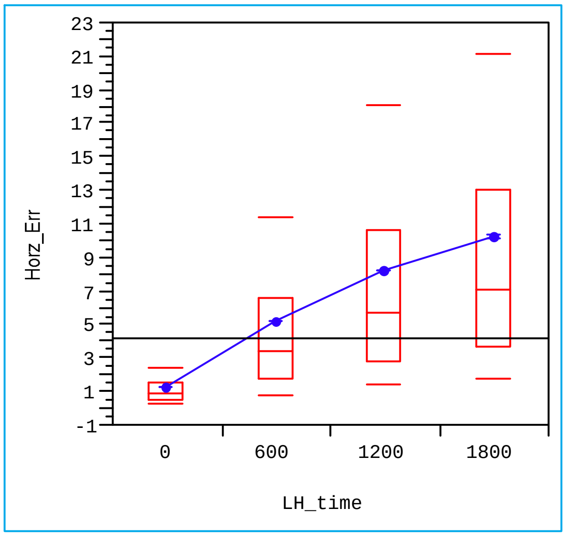

Figure 3.3-3: Quantile / Mean Comparison of Horizontal Error Vs. LH

In summary, the mean horizontal error is statistically significant at the look ahead times of zero, 600, 1200, and 1800 seconds. Referring to Figure 3.3-3, the sample means are also increasing as the look ahead time (LH) increases, ranging from a sample mean of 1.2 nautical miles at look ahead zero to 10.2 at 1800 seconds (i.e. 30 minutes). The mean of all observations is drawn as a horizontal line across the entire plot. The median is also increasing from 0.96 nautical miles at zero look ahead time to 7.1 at 1800 seconds. The horizontal lines in Figure 3.3-3's boxes correspond to the 10, 25, 50, 75, and 90 percentiles of the distribution of the sampled horizontal errors, respectively[7]. Tested statistically with the Levene Test earlier, the box ranges illustrate that the spread of the horizontal error is also increasing as the look ahead time increases.

The analysis continues by examining the lateral, longitudinal, and vertical errors using the same methods described for the horizontal error. The results are summarized in Table 3.3-6 and the means comparisons of the lateral, longitudinal and vertical errors are shown in Figures 3.3-4 through 3.3-6. The descriptive statistics of the absolute values of the four errors are tabulated in Appendix A.1.

[7] The percentiles illustrated in the Figure 3.3-3 as horizontal lines and box ends are described in detail in Appendix A.0.

Table 3.3-6: Statistical Results LH 0-30 minutes for All Altitudes

Error Type	Levene Test	Welch Test	Tukey-Kramer[8]	Observations
Horizontal	Yes	Yes	Yes – all	Mean and variance increases as look ahead time (LH) increases. Means range from 1.2 to 10.2 nautical miles (nm).
Lateral	Yes	Yes	Yes-3of6	Mean at LH 0 different from others. Mean and variance increase as LH increases. Means range from -0.02 to -0.22 nm.
Longitudinal	Yes	Yes	Yes – 5of6	Both mean and variance different. Only means at LH 1200 versus 1800 not different. Means increase in value as LH increases, ranging from –0.02 to 0.88 nm.
Vertical	Yes	Yes	Yes –all	Mean ranges from 49 to –327 feet. Mean (becomes more negative) and variance increase as LH increases.

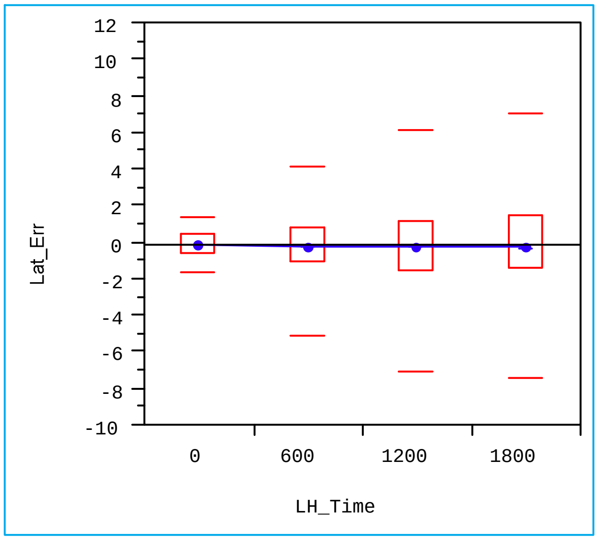

Figure 3.3-4: Quantile / Mean Comparison of Lateral Error Vs. LH

[8] In this table, "yes" means test provides evidence to reject hypothesis that means or variances are equal. "Yes-all" means Tukey-Kramer found all pairs of means not equal, and "Yes-1of6" means it found only 1 pair of means not equal in 6 combinations of pairwise comparisons.

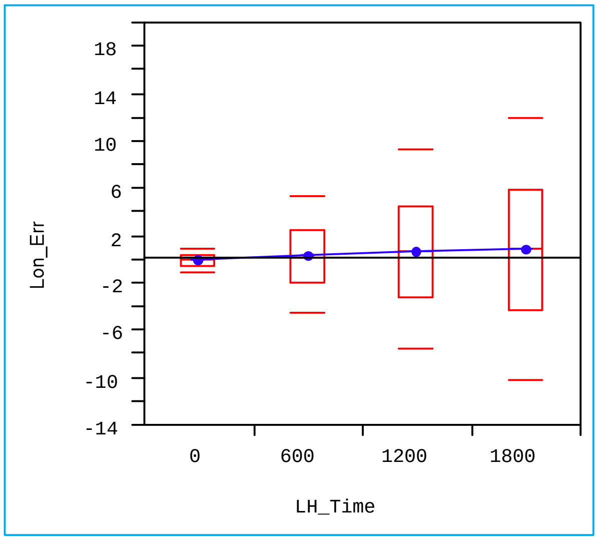

Figure 3.3-5: Quantile / Mean Comparison of Longitudinal Error Vs. LH

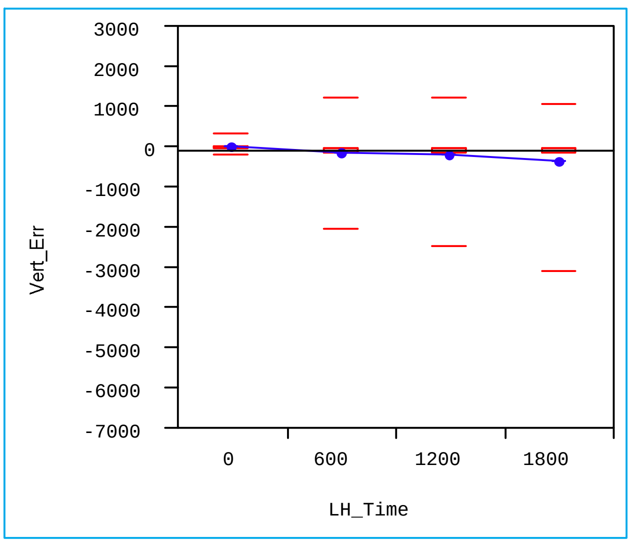

Figure 3.3-6: Quantile / Mean Comparison of Vertical Error Vs. LH

3.3.1.2 Samples at altitudes above 18,000 feet

For samples at altitudes above 18,000 feet only, the results are summarized in Table 3.3-7. The detailed histograms and statistical tables are located in Appendix A.1.

Table 3.3-7: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	Yes	Yes	Yes – all	Mean and variance increases as LH increases. Means range from 1.1 to 10.6 nm and standard deviation ranges from 0.94 to 11.5 nm.
Lateral	Yes	Yes	Yes-3of6	Only LH 0 different from others. Variance increases as LH increases. Means range from –0.02 to -0.44 nm.
Longitudinal	Yes	Yes	Yes – 5of6	Mean LH 1200 versus 1800 not different. Mean and variance increase with LH.
Vertical	Yes	Yes	Yes – 5of6	Mean ranges from 39 to -180 feet. Variance increases with LH. T-K Test shows no difference between means at 0 and 600 seconds LH.

3.3.1.3 Discussion of the effect of look ahead time

In general, look ahead time does have a significant effect on each sample mean and increases as the look ahead time increases. For horizontal error, the sample means increase over 10 nautical miles from zero to 1800 seconds (i.e. 30 minutes) look ahead time. Since lateral and longitudinal errors are exact orthogonal components of the horizontal error, it is interesting to note that the dominant source of the increase in horizontal error with look ahead time is the longitudinal error. Longitudinal error increases around one nautical mile with look ahead time zero to 30 minutes, while the absolute longitudinal error does increase around seven nautical miles. The lateral error increases by around a 0.25 nautical mile with look ahead time, and its absolute error increases by around four nautical miles. Statistically the lateral error only shows a difference between look ahead zero and the others, while longitudinal shows a difference in practically all look ahead times except between 1200 and 1800 seconds. Therefore, most of the error affecting an increase in the horizontal dimension as look ahead time increases is dominated by the longitudinal component.

Another aspect of the longitudinal error is the direction of the increase as look ahead time increases. On average, longitudinal error becomes more positive as look ahead increases. The aircraft on average are getting ahead of the prediction or conversely the predictions are getting behind the aircraft. The specific reasons for this will have to be left for future study but could be related to anything from URET's aircraft modeling parameters to weather profiles of the particular day analyzed.

The vertical error also shows a significant difference between sample means, but the mean differences like the lateral error are relatively small, ranging around 300 to 400 feet for all altitudes and around 200 feet for samples above 18,000 feet. For the vertical error, the sample means may be relatively small, but the variance increases dramatically with a standard deviation ranging from around 600 to 2300 feet. In other words, the central tendency of the vertical error may not change dramatically, but the spread increases significantly as look ahead time increases.

In general, the variance increases significantly for all the error variables in both horizontal and vertical dimensions. For horizontal error, the standard deviation increases over nine nautical miles from zero to 1800 look ahead time. This range of nine nautical miles holds true for lateral and longitudinal errors as well. The spread of the errors increases as the look ahead time increases.

The differences between the trajectory prediction errors from samples at all altitudes versus above 18,000 feet are small, and they lead to the same conclusions about the distributions.

3.3.2 Analysis of Flight Type on Trajectory Accuracy

Flight type is determined by examining the origin and destination airports in a flight plan. The flight type includes four possible levels referred to as overflight, departure, arrival, and internal. Overflight is an aircraft whose origin and destination airports are outside the particular center's airspace, ZID in this case. Departures leave an airport inside the center, and arrivals land at an airport inside the center. The internals include flights that have both origin and destination airports inside the center.

The analysis that follows examines whether the means of the trajectory prediction errors of the different flight types are significantly different at the four look ahead times of 0, 600, 1200, and 1800 seconds. This analysis focuses on these four look ahead times and flight types against the signed lateral, longitudinal, vertical, and horizontal errors. Appendix A.1 contains a more complete set of look ahead times and also includes the descriptive statistics on the unsigned or absolute values of the errors. Figures 3.3-7 through 3.3-10 plot the means as a function of look ahead time (LH) where OVR denotes overflights, ARR denotes arrivals, DEP denotes departures, and INR denotes internals.

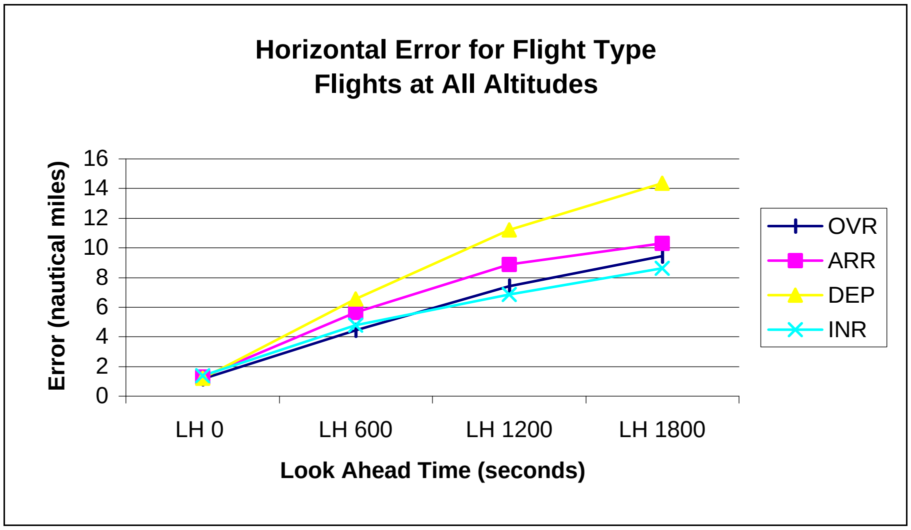

Figure 3.3-7: Sample Means for Horizontal Error per Flight Type and LH

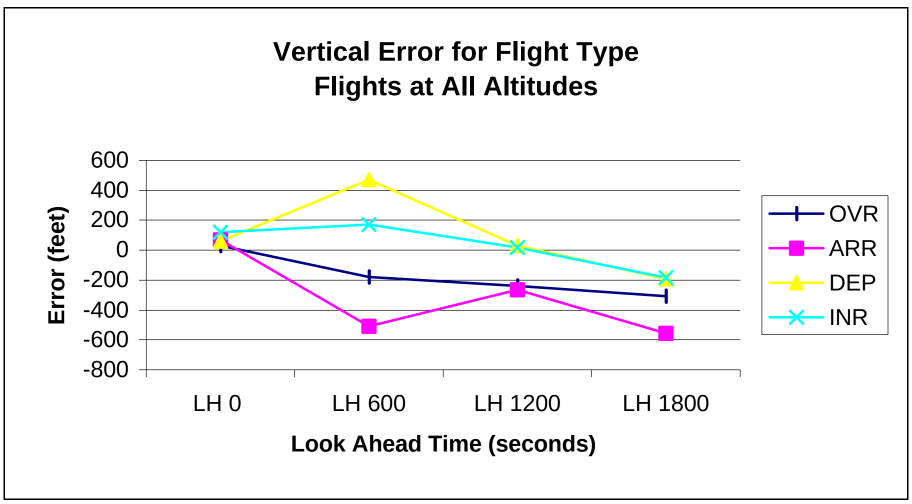

Figure 3.3-8: Sample Means for Vertical Error per Flight Type and LH

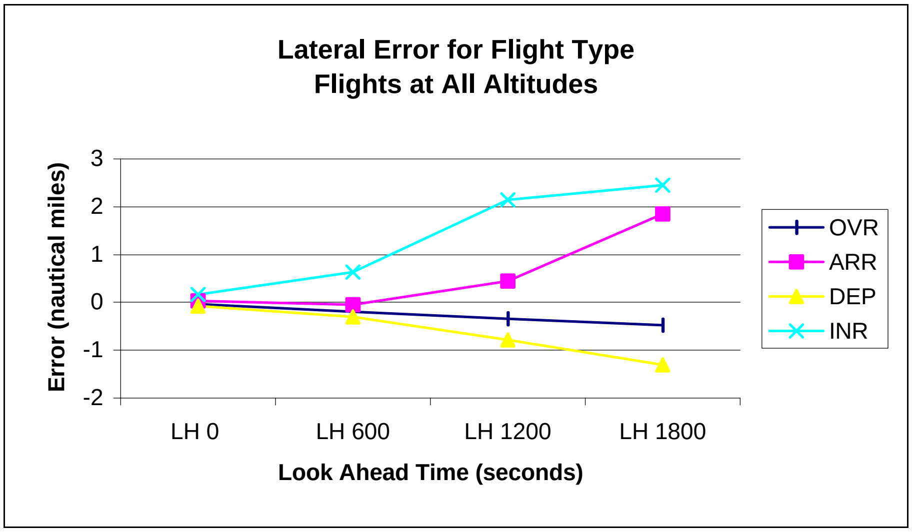

Figure 3.3-9: Sample Means for Lateral Error per Flight Type and LH

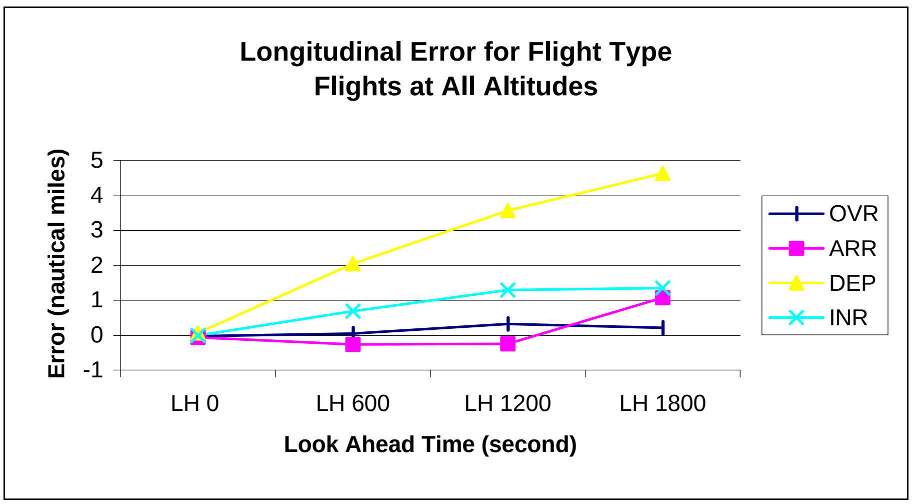

Figure 3.3-10: Sample Means for Longitudinal Error per Flight Type and LH

3.3.2.1 Samples at all altitudes

Statistical results for all altitudes are summarized in Table 3.3-8. The detailed histograms and statistical tables are located in Appendix A.1.

Table 3.3-8: Statistical Results LH 0-30 minutes at All Altitudes

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-5of6	T-K Test shows arrivals and internals are not significantly different.
Lateral	0	Yes	Yes	Yes-5of6	Only overflights versus departures are not different.
Long.	0	Yes	Yes	Yes-3of6	Only internals are not significantly different from the others.
Vertical	0	Yes	Yes	Yes-3of6	Only overflights are different than the other three flight types. Overflights have less vertical error with a sample mean of 32 feet compared to a range of 61-121 feet.
Horizontal	600	Yes	Yes	Yes-5of6	T-K shows overflights and internals are not significantly different.
Lateral	600	Yes	Yes	Yes-2of6	Only internals versus either overflights or departures are significantly different.
Long.	600	Yes	Yes	Yes-5of6	Only internals versus overflights are not different.
Vertical	600	Yes	Yes	Yes-all	Although all the means are different, arrivals and departures are around 500 feet in error on average and overflights and internals are around 200 feet.
Horizontal	1200	Yes	Yes	Yes-5of6	T-K shows overflights and internals are not significantly different.
Lateral	1200	Yes	Yes	Yes-5of6	Only overflights versus departures are not significantly different.
Long.	1200	Yes	Yes	Yes-4of6	Only internals versus either overflights or arrivals are not significantly different.
Vertical	1200	Yes	Yes	Yes-2of6	Departures versus overflights or arrivals are significantly different.
Horizontal	1800	Yes	Yes	Yes-3of6	T-K shows only departures are significantly different to the other types.
Lateral	1800	Yes	Yes	Yes-3of6	Departures versus either arrivals or internals and arrivals versus overflights are significantly different.
Long.	1800	Yes	Yes	Yes-2of6	Departures versus arrivals and overflights are significantly different.
Vertical	1800	Yes	Yes	Yes-2of6	All means negative ranging from 200 to 600 feet error. T-K shows arrivals versus overflights and departures are different.

3.3.2.2 Samples at altitudes above 18,000 feet

Statistical results for altitudes above 18,000 feet are summarized in Table 3.3-9. The detailed histograms and statistical tables are located in Appendix A.1.

Table 3.3-9: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-4of6	T-K Test shows internals versus overflights and arrivals are not different.
Lateral	0	Yes	Yes	Yes-1of6	Only arrivals versus departures are significantly different.
Long.	0	Yes	Yes	Yes-3of6	Only internals are not different from the other flight types.
Vertical	0	Yes	Yes	Yes-4of6	T-K shows departures versus overflights and arrivals are significantly different. Overflights and departures have less error with around 32 feet on average.
Horizontal	600	Yes	Yes	Yes-3of6	Only internals vs. others are not different.
Lateral	600	Yes	Yes	Yes-1of6	Only overflights versus departures are significantly different.
Long.	600	Yes	Yes	Yes-5of6	Only internals versus departures are not different.
Vertical	600	Yes	Yes	Yes-all	All means are different ranging from around -168 to 3700 feet.
Horizontal	1200	Yes	Yes	Yes-3of6	Only internals vs. others are not different, based on one sample so inconclusive.
Lateral	1200	Yes	Yes	Yes-3of6	All are different except internals which are based on one sample.
Long.	1200	Yes	Yes	Yes-2of6	Departures versus overflights and arrivals are different. Only one sample for internals.
Vertical	1200	Yes	Yes	Yes-all	All means are significantly different, but internals inconclusive with one sample.
Horizontal	1800	Yes	Yes	Yes-2of3	No internal samples. Departures differ from overflights and arrivals.
Lateral	1800	Yes	Yes	Yes-All	No internal samples. All means and variance different.
Long.	1800	Yes	Yes	Yes-2of3	No internal samples. Only overflights and arrivals are not different.
Vertical	1800	Yes	Yes	Yes-2of3	No internal samples. Arrivals differ from overflights and departures.

3.3.2.3 Discussion of the effect of flight type

In general, flight type did have a significant effect on the performance of the trajectory predictions but not nearly as much as the look ahead time. In general, overflights performed the best at the lower look ahead times for all samples, but internals and overflights did not have significant differences at the larger look ahead times for all altitude samples. Any conclusions on internals for the samples above 18,000 feet are inconclusive since the sample sizes were small or nonexistent. For horizontal error, departures seem to have the largest error, ranging from 1.2 to 14.4 nautical miles, as look ahead time increases. For vertical error, the same is true for arrivals. That is, for arrivals the vertical error increases as look ahead time increases the most from around 60 to –550 feet on average.

There were relatively small sample sizes for internals at the larger look ahead times. The samples are taken along a trajectory for a look ahead time window up to 30 minutes (i.e. 1800 seconds), but the internals have much shorter flights on average. The internals have an average track life of around 22 minutes, compared to the other flight types which have an average track life of around 35 minutes.

3.3.3 Analysis of Horizontal Phase of Flight on Trajectory Accuracy

Horizontal phase of flight is calculated for each HCS track report and extracted for the trajectory accuracy measurements. This factor is categorized into two levels: straight or turn. The PHASE_D program that detects turns, described in Section 2.4.6.1, had its parameters set to protect against noise in the track data. As a result, rapid turns are detected but shallow turns may be missed. A turn is determined by a nine degree angle (or greater) generated by the two segments drawn from the previous position to the current position and the current position to the next position report.

The analysis that follows examines whether the mean of the trajectory prediction error at the two horizontal phases of flight are significantly different statistically at the four look ahead times of zero, 600, 1200, and 1800 seconds. This analysis will focus on these four look ahead times and two phases of flight against the signed lateral, longitudinal, vertical, and horizontal errors. Appendix A.1 contains a more complete set of look ahead times and also includes the descriptive statistics on the unsigned or absolute values of the errors. Figures 3.3-11 to 3.3-14 plot the means for each horizontal phase of flight as a function of look ahead time (LH), where STR denotes straight and TRN denotes turning.

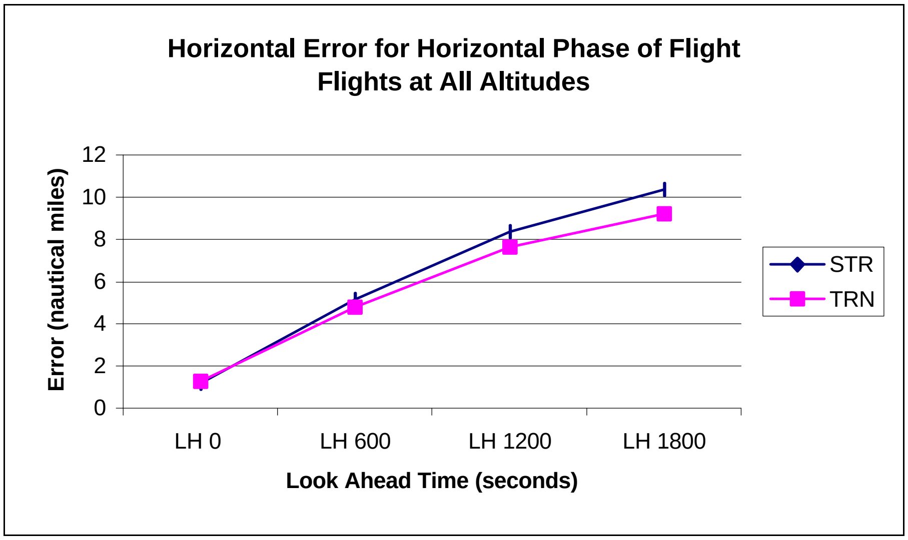

Figure 3.3-11: Sample Means for Horizontal Error per Horizontal Phase of Flight and LH

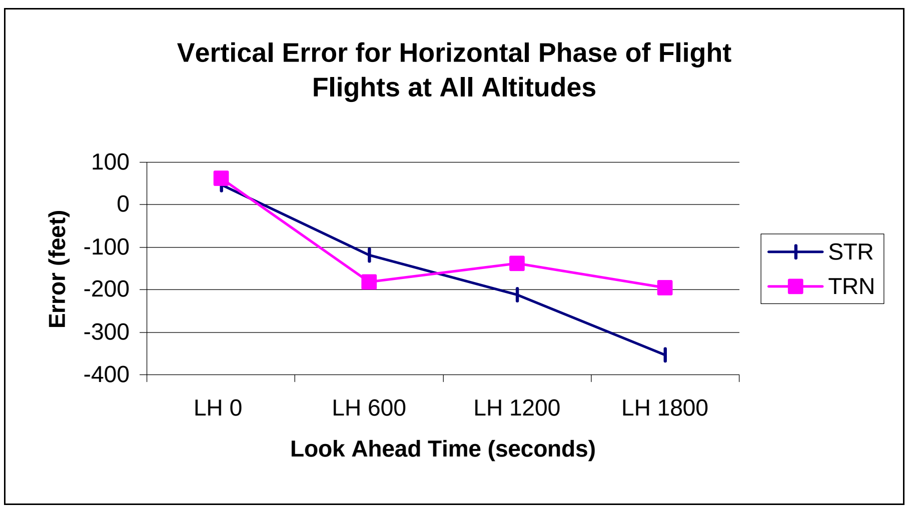

Figure 3.3-12: Sample Means for Vertical Error per Horizontal Phase of Flight and LH

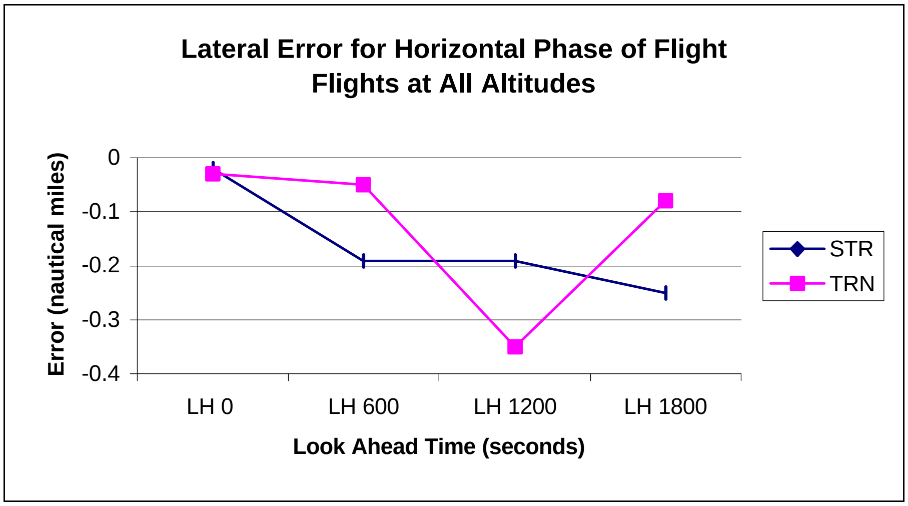

Figure 3.3-13: Sample Means for Lateral Error per Horizontal Phase of Flight and LH

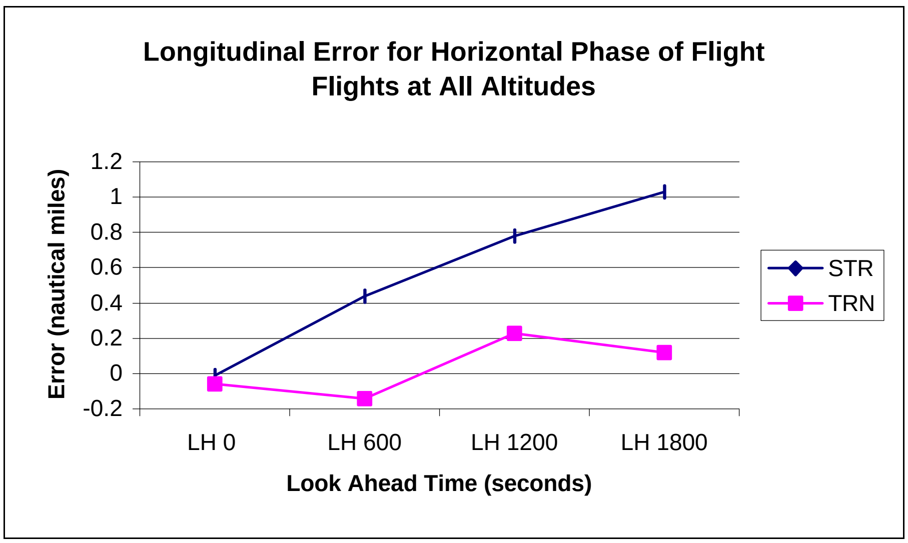

Figure 3.3-14: Sample Means for Longitudinal Error per Horizontal Phase of Flight and LH

3.3.3.1 Samples at all altitudes

The results for all altitudes are summarized in Table 3.3-10. The detailed histograms and statistical tables are located in Appendix A.1.

Table 3.3-10: Statistical Results LH 0-30 minutes at All Altitudes

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes	Both mean and variance are significantly different. The means are around 600 feet different.
Lateral	0	Yes	No	No	Only variance is significantly different.
Long.	0	Yes	Yes	Yes	Both mean (around 300 feet) and variance are significantly different.
Vertical	0	Yes	No	No	Only variance is significantly different.
Horizontal	600	Yes	Yes	Yes	Both mean (around 900 feet) and variance are significantly different.
Lateral	600	Yes	No	No	Only variance is significantly different.
Long.	600	No	Yes	Yes	Means are different, around 0.6 nm.
Vertical	600	Yes	No	No	Only variance is significantly different.
Horizontal	1200	Yes	Yes	Yes	Both mean and variance are significantly different. The means differ around 1 nautical mile.
Lateral	1200	Yes	No	No	Only variance is significantly different.
Long.	1200	No	Yes	Yes	Means are significantly different. The means differ around 0.5 nautical mile.
Vertical	1200	No	No	No	Do not differ statistically.
Horizontal	1800	Yes	Yes	Yes	Both mean and variance are significantly different. The means differ around 1.2 nautical miles.
Lateral	1800	Yes	No	No	Only variance is significantly different.
Long.	1800	No	Yes	Yes	Means are significantly different. The means differ 0.9 nm.
Vertical	1800	No	Yes	Yes	Means are significantly different. The means differ around 160 feet.

3.3.3.2 Samples at altitudes above 18,000 feet

The results are summarized in Table 3.3-11. The detailed histograms and statistical tables are located in Appendix A.1.

Table 3.3-11: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	No	No	Only variance is significantly different.
Lateral	0	No	No	No	Do not differ statistically.
Long.	0	No	No	No	Do not differ statistically.
Vertical	0	No	No	No	Do not differ statistically.
Horizontal	600	No	No	No	Do not differ statistically.
Lateral	600	No	No	No	Do not differ statistically.
Long.	600	No	Yes	Yes	Means differ around a 0.6 nm.
Vertical	600	No	No	No	Do not differ statistically.
Horizontal	1200	Yes	Yes	Yes	Both mean and variance are significantly different. The means differ around 1 nautical mile.
Lateral	1200	Yes	No	No	Only variance is significantly different.
Long.	1200	No	No	No	Do not differ statistically.
Vertical	1200	No	No	No	Do not differ statistically.
Horizontal	1800	Yes	Yes	Yes	Both mean and variance are significantly different. The means differ around 1.3 nautical miles.
Lateral	1800	Yes	No	No	Only variance is significantly different.
Long.	1800	No	Yes	Yes	Means are significantly different. The means differ around 1.3 nm.
Vertical	1800	No	Yes	Yes	Means are significantly different. The means differ around 230 feet.

3.3.3.3 Discussion of the effect of Horizontal Phase of Flight

In general, the horizontal phase of flight, i.e. whether an aircraft is turning or on a straight path, had a significant effect on the horizontal prediction error and longitudinal error only for the all altitude samples. The magnitude of these differences between the means was rather small, approximately 0.1 to 1.2 nautical miles from zero to 1800 seconds look ahead time, respectively. The only other pattern of significant differences between means was the vertical error at 1800 seconds look ahead time, however the differences were very small, at around 150 feet. The results suggest that horizontal phase of flight has only a minor impact on the trajectory performance. There has also been some discussion on the need for analysis a small distance before and after the actual turn. The technique currently used for determining an aircraft is turning is not sufficiently robust in filtering out the noise of the HCS track reports nor can it examine the straight path around the turn. As a result, the statistical analysis of the effect of turns should be interpreted advisedly and the algorithm will be revisited in the future.

3.3.4 Analysis of Vertical Phase of Flight on Trajectory Accuracy

Similar to horizontal phase of flight, vertical phase of flight is calculated for each interpolated HCS track report and extracted for the trajectory accuracy measurements. Vertical phase of flight is categorized into three categories: level, ascending, or descending. The track points are only labeled as climbing or descending for reasonably large climbs and descents to protect against noise in the position data, but this also prevents detection of low rate climbs and descents (i.e. smaller than 900 feet per minute). A climb or descent is determined by calculating the difference in altitude between the current interpolated track position and the next track position. If the absolute difference is less than 150 feet, the current position of the aircraft is considered in level flight, otherwise the aircraft is in a climb or descent depending on the direction up or down. Since the track positions are interpolated at 10 second intervals, the required gradient for the climbing or descending aircraft is greater than or equal to 15 feet per second or 900 feet per minute. The phase of flight algorithm is described in detail in Section 2.4.6.

The analysis that follows examines whether the mean of the trajectory prediction error at the three vertical phases of flight are significantly different statistically at the four look ahead times of zero, 600, 1200, and 1800 seconds. This analysis focuses on these four look ahead times and three phases of flight against the signed lateral, longitudinal, vertical, and horizontal errors. Appendix A.1 contains a more complete set of look ahead times and also includes the descriptive statistics on the unsigned or absolute values of the errors. Figures 3.3-15 to 3.3-18 plot the means for each vertical phase of flight as a function of look ahead time (LH), where LEV denotes level flight, ASC denotes ascending and DES denotes descending.

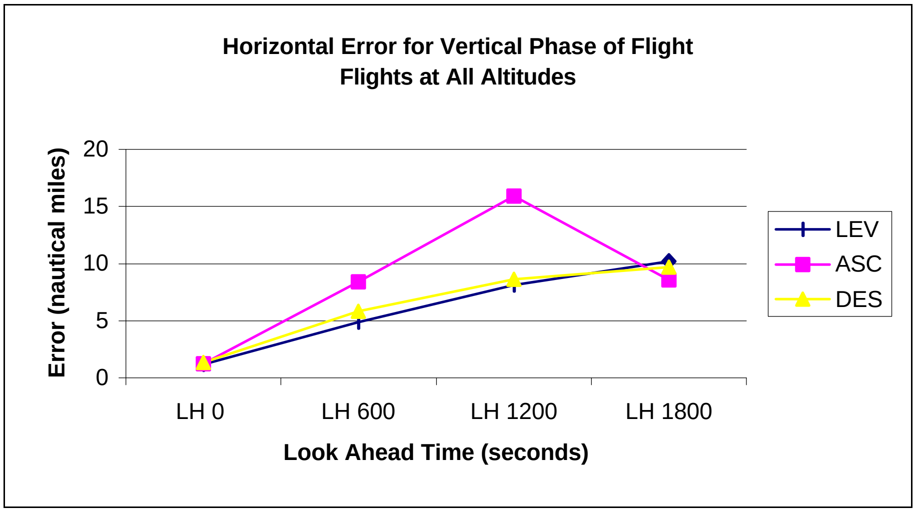

Figure 3.3-15: Sample Means for Horizontal Error per Vertical Phase of Flight and LH

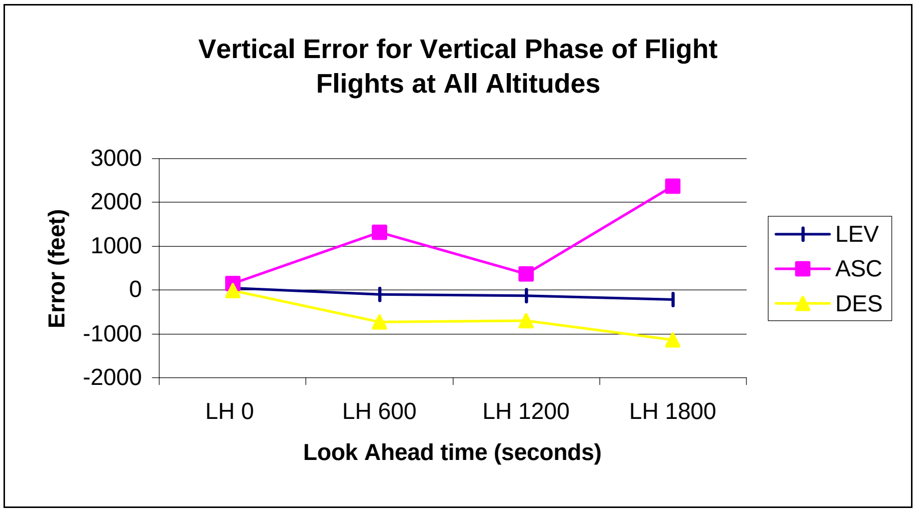

Figure 3.3-16: Sample Means for Vertical Error per Vertical Phase of Flight and LH

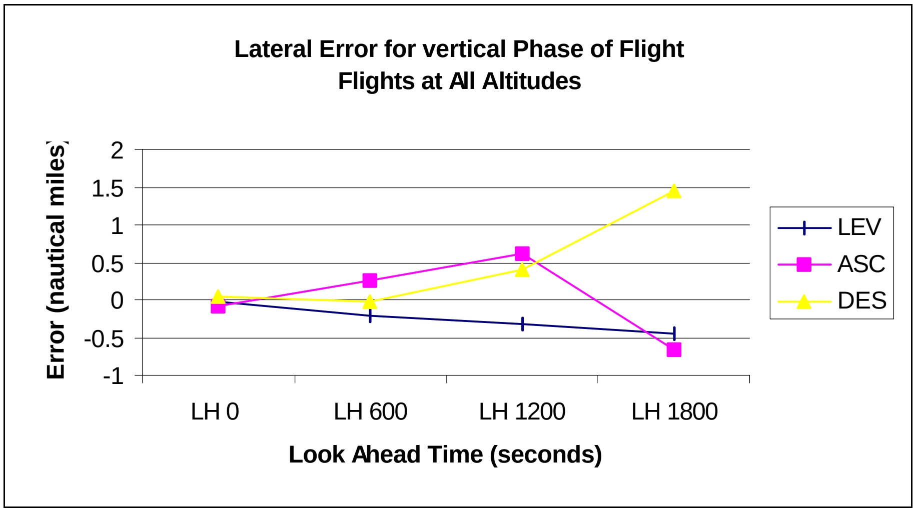

Figure 3.3-17: Sample Means for Lateral Error per Vertical Phase of Flight and LH

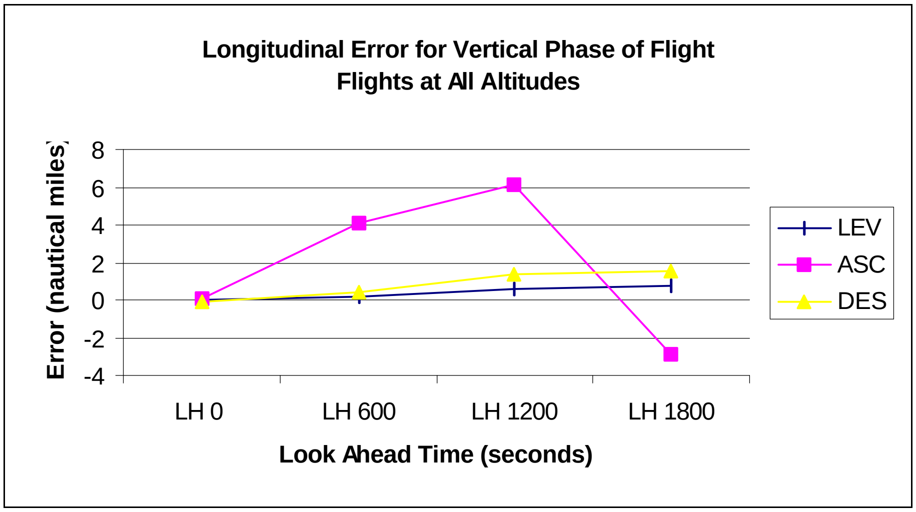

Figure 3.3-18: Sample Means for Longitudinal Error per Vertical Phase of Flight and LH

3.3.4.1 Samples at all altitudes

The results are summarized in Table 3.3-12. The detailed histograms and statistical tables are located in Appendix A.1.

Table 3.3-12: Statistical Results LH 0-30 minutes at All Altitudes

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-2of3	Only level versus ascent not different, but others around a maximum of 1000 feet different.
Lateral	0	Yes	Yes	Yes-all	Both mean and variance are significantly different.
Long.	0	Yes	Yes	Yes-all	Both mean and variance are significantly different.
Vertical	0	Yes	Yes	Yes-all	Both mean (around 160 feet) and variance are significantly different.
Horizontal	600	Yes	Yes	Yes-all	Both mean (by as much as 3.6 nm) and variance are significantly different.
Lateral	600	Yes	Yes	Yes-1of3	Only ascent versus level differ.
Long.	600	Yes	Yes	Yes-all	Both mean (by as much as 3.9 nm) and variance are significantly different.
Vertical	600	Yes	Yes	Yes-all	Both mean (by as much as 2000 feet) and variance are significantly different.
Horizontal	1200	Yes	Yes	Yes-2of3	Only level versus descent not different, and others differ by as much as 7.75 nm.
Lateral	1200	No	Yes	Yes-1of3	Only means descent versus level are significantly different.
Long.	1200	Yes	Yes	Yes-all	Mean (by as much as 5.5 nm) and variance are significantly different.
Vertical	1200	Yes	Yes	Yes-2of3	Both mean and variance are significantly different, except level versus ascent. The means differ by as much as 1100 feet.
Horizontal	1800	Yes	No	No	Only variance is significantly different. Inconclusive on ascents, only 11 samples.
Lateral	1800	No	Yes	Yes1of3	Only mean of descent versus level different. Inconclusive on ascents, only 11 samples.
Long.	1800	No	No	No	Do not differ statistically. Inconclusive on ascents, only 11 samples.
Vertical	1800	Yes	Yes	Yes	Means are significantly different. The means differ by as much as 3500 feet. Inconclusive on ascents, only 11 samples.

3.3.4.2 Samples at altitudes above 18,000 feet

The results are summarized in Table 3.3-13. The detailed histograms and statistical tables are located in Appendix A.1.

Table 3.3-13: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-2of3	Only level versus ascent not different, but others around a maximum of 600 feet different.
Lateral	0	Yes	Yes	No	Tukey-Kramer shows no difference in means but has less power than Welch Test which had a p-value of 0.48.
Long.	0	Yes	Yes	Yes-2of3	Only descent versus level means are not significantly different
Vertical	0	Yes	Yes	Yes-all	Both mean (around 220 feet) and variance are significantly different.
Horizontal	600	Yes	Yes	Yes-all	Both mean (by as much as 3.5 nm) and variance are significantly different.
Lateral	600	Yes	No	No	Only variance is significantly different.
Long.	600	Yes	Yes	Yes-2of6	Only descent versus level means are not significantly different.
Vertical	600	Yes	Yes	Yes-all	Both mean (by as much as 1600 feet) and variance are significantly different.
Horizontal	1200	Yes	Yes	Yes-2of3	Only level versus descent not different, and others differ by as much as 7 nm.
Lateral	1200	No	Yes	Yes-1of3	Only descent versus level are significantly different.
Long.	1200	Yes	Yes	Yes-2of3	Only descent versus level not different, and others differ by as much as 8.26 nm.
Vertical	1200	Yes	Yes	Yes-2of3	Except level versus ascent means, both mean and variance are different. The means differ by as much as 970 feet.
Horizontal	1800	Yes	Yes	Yes-1of3	Only descent versus level are different, around 2 nautical miles. Inconclusive on ascents, only 10 samples.
Lateral	1800	Yes	Yes	Yes-1of3	Only descent versus level means are different, around 1.75 nautical miles. Inconclusive on ascents, only 10 samples.
Long.	1800	Yes	No	No	Only variance is significantly different. Inconclusive on ascents, only 10 samples.
Vertical	1800	Yes	Yes	Yes-all	Means differ by as much as 3300 feet. Inconclusive on ascents, only 10 samples.

3.3.4.3 Discussion of the effect of Vertical Phase of Flight

In general for both horizontal and vertical dimensions, level flight has the smallest mean and variance error, while ascending flight has the largest as look ahead time increases. At a look ahead time of zero, both ascent and level are not significantly different, but at look ahead time of 1800 not much can be drawn on ascending flight from these samples because around 10 samples were available. In practically all cases, the variance was significantly different. Also as the look ahead time increases, the standard deviation increases and the difference in standard deviation between levels increases. For example, for vertical error at look ahead time zero seconds, the standard deviation ranges from around 620 feet to 940 feet, but at look ahead time 1200 seconds the standard deviation ranges from around 1860 feet to 3200 feet.

4. CTAS Study Results and Observations

The results and observations presented in this section are based on the analysis of seven hours of data recorded at the Fort Worth ARTCC (ZFW). Specific information describing the scenario is presented in Section 4.1. Section 4.2 provides detailed information about one aircraft flight in the study in order to demonstrate the study's methodology, and Section 4.3 presents the results of the application of the trajectory accuracy metrics to CTAS.

4.1 Scenario Description

Figure 4.1-1 provides a data flow diagram logically describing the data files and processes used to obtain the flight plan, track, and trajectory data used for the CTAS analysis. For this study, data was collected from the CTAS installation at ZFW. A recording was made of the HCS flight plans, flight plan amendments, and track messages sent to CTAS over a seven hour period on January 5, 1999. The weather data for the same time period was also recorded.

NASA Ames Research Center provided the ZFW data to ACT-250 in file called *ZFW_010599.cm_sim*. This file was used as input to a playback run through a developmental version of CTAS also provided by NASA Ames. This version of CTAS, called *daisy_view*, was modified by ACT-250 to provide trajectories in its output file. These trajectories consist of 31 points, each point separated in time by 65 seconds. As a result, all of the CTAS trajectories were 1950 seconds or less in length. This output file is identified as *baseline.cm_sim* in Figure 4.1-1. The CTAS Parser Program (CPP) used the *baseline.cm_sim* file to create three files: the *fp.dat* file, containing flight plan data; the *track.dat* file, containing track data; and the *traj_file.dat* file, containing trajectory data. The *fp.dat* file was then concatenated with the *track.dat* file to create an ASCII file called *sn010599.dat,* containing CTAS field data, that has the same format as the *sn022798.dat* described for URET field data in Section 3.1. The *sn010599.dat* file was then used as input to the Flight Plan and Track Data Processing described in Section 2.4.1. The *traj_file.dat* file has the same format as its URET counter part described in Section 3.1 and was used as input to the Trajectory Data Processing described in Section 2.5. The formats of the *sn010599.dat* and *traj_file.dat* files are described in WJHTC/ACT-250, 1998.

Tables 4.1-1 and 4.1-2 summarize the characteristics of the airspace and the aircraft flights through the airspace, respectively, for the subject scenario.

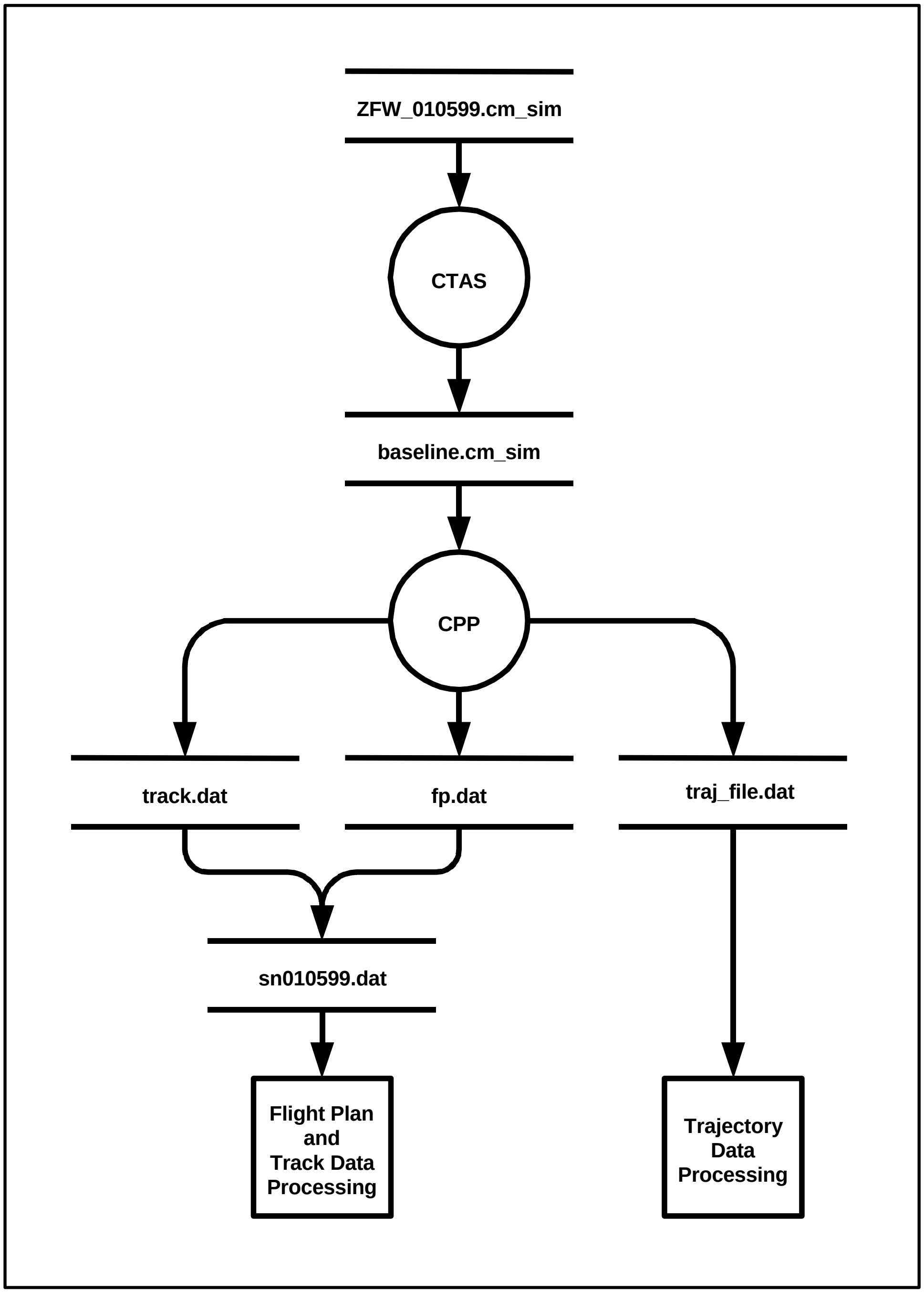

Figure 4.1-1: CTAS Data Sources

4.1.1 Airspace Definition

Table 4.1-1 summarizes the spatial and time boundaries of the ZFW data sample used.

Table 4.1-1: CTAS Scenario - Airspace

Airspace	Fort Worth (ZFW)
Altitude	0 to 60,000 feet
Horizontal boundaries	Defined by the high altitude sectors
Date	January 5, 1999
Start time	18:39:35 UTC (12:40 p.m. local time)
End time	01:43:26 UTC (7:43 p.m. local time)
Duration	07:03:51 or 25,431 seconds

4.1.2 Aircraft Counts

Table 4.1-2 gives the counts of aircraft flights in the sample of air traffic analyzed.

Table 4.1-2: CTAS Scenario – Aircraft Counts

Total number in sample (IFR)	2592
Number excluded	297 (11.5 %)
Number processed	2295 (88.5 % of total)
Number of airliners	1699
Number of General Aviation aircraft	596
Number of jet types in the top 20 aircraft	15
Number of turboprop types in the top 20 aircraft	4
Number of piston types in the top 20 aircraft	1
Average length of track supplied by HCS	37.6 minutes, 2253 seconds, or 189 position reports
Number of overflights	604 (26.3 %)
Number of departures	683 (29.8 %)
Number of arrivals	719 (31.3 %)
Number of internal flights	289 (12.6 %)

4.1.3 Excluded Flights

In measuring the accuracy of track predictions, the true positions of the aircraft are assumed to be the positions reported by the HCS. For some aircraft, it is clear that the HCS reported positions are not correct. Track processing algorithms (in the RDTRACKS program) were used to correct the position data where possible, as described in Section 2.4.3. When it was not possible to correct the data, the individual track reports and in some cases entire flights were deleted from the scenario being examined. Statistics were collected on an aircraft flight only if both a track and a set of predicted trajectories were available. For this analysis of CTAS, there were three categories of excluded aircraft, totaling 297 flights that were deleted from the original set of 2592 IFR flights (a reduction of 11.5 %).

4.1.3.1 Military Flights

Since it is often not possible from flight plan data to accurately predict the flight paths of military flights, which usually are doing either gunnery practice or aerial re-fueling maneuvers, military flights were excluded from the analysis. This was done by selecting out all of the flights which had a call sign containing more than three leading alphabetic characters (e.g., ANVIL, CODER, RACER, SABER, STEEL). Although this is not an exact definition of military aircraft, it was considered to be sufficient for this study. 99 military flights were excluded.

4.1.3.2 Non-initialized Flights

As discussed in Section 2.4, sometimes the HCS processing algorithms are unable to establish a consistent track for the aircraft. Ten of these flights were excluded.

4.1.3.3 Uncertain Position Flights

The processing of the HCS track data requires correcting some of the track reports which are clearly in error. For example, as discussed in Section 2.4.3, sometimes the same XY coordinates are repeated even though the aircraft has moved between the radar reports. In some cases the corrected track reports are substantially different from the original aircraft positions reported by the HCS. This situation implies that we, the experimenters, do not know the true position of the aircraft. Flights having a corrected position report substantially different from the original position report were deleted (188 of these flights were excluded).

4.1.4 Truncated Flights

Often in the HCS track reports several tracks reports are missing or have bad data. If a gap in the track data is short, the missing track reports can be replaced by interpolation. If the gap is large, the position of the aircraft during the gap is unknown. When a large gap in the track data occurs, the track after the gap is discarded. Of the 441,557 radar track position reports, 14,333 or 3.2 % of the radar track position reports were discarded by truncating the tracks after missing or bad data.

Measurements of trajectory prediction errors were made on aircraft either already in the ZFW airspace or approaching the ZFW airspace and about to be in the ZFW airspace. Measurements were not made on aircraft after they left the ZFW airspace. That is, no measurements were made on the portions of the tracks outside ZFW when the aircraft were flying away from the ZFW airspace. 12.6 % of the interpolated track reports were not used for this reason.

4.1.5 Aircraft Mix

The majority of the aircraft in the study are commercial airliners. The top 10 aircraft types account for 1310 of the 2295 flights, or 57.1 % of the total; the top 20 aircraft account for 1632 of the 2295 flights, or 71.1 % of the total. A histogram depicting the frequency of occurrence of the top 20 aircraft is provided in Figure 4.1-2. The aircraft are identified by their FAA type designators. Of the top 20 aircraft, 15 are jets, four are turboprops, and one is a piston-powered aircraft. Table 4.1-3 lists the aircraft manufacturers and model names of the top 10 aircraft. All of the top 10 aircraft are jets except for the Saab & Fairchild 340 which is a turboprop.

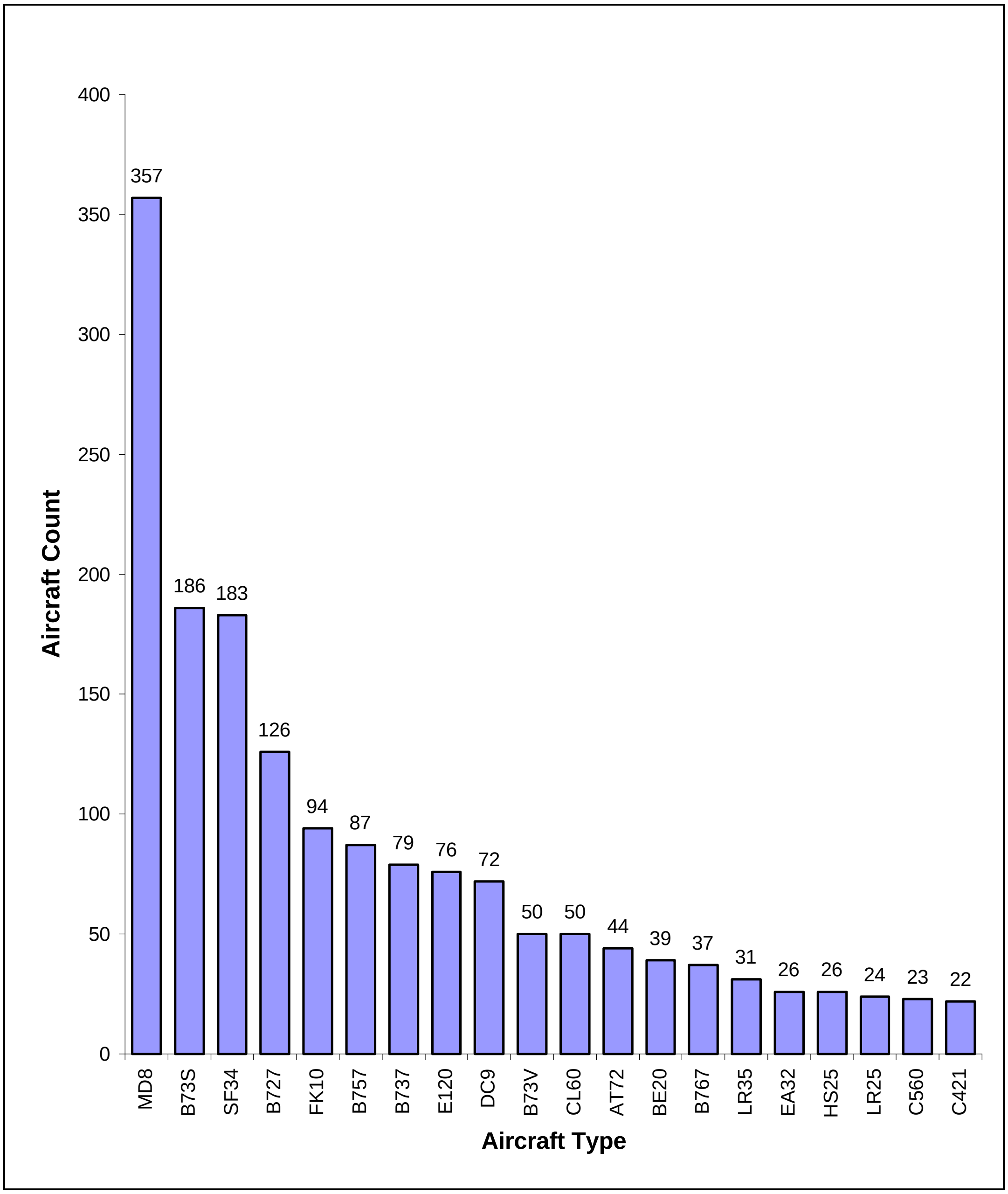

Figure 4.1-2: Top 20 Aircraft Frequency Histogram - ZFW Data

Table 4.1-3: CTAS Scenario Aircraft

RANK	FAA TYPE IDENTIFIER	MANUFACTURER / MODEL	NUMBER OF FLIGHTS	PERCENTAGE OF FLIGHTS
1	MD8	McDonnell-Douglas MD-80	357	15.56 %
2	B73S	Boeing 737 – 300/500	186	8.10 %
3	SF34	Saab & Fairchild 340	183	7.97 %
4	B727	Boeing 727	126	5.49 %
5	FK10	Fokker 100	94	4.10 %
6	B757	Boeing 757	87	3.79 %
7	B737	Boeing 737 – 200	79	3.44 %
8	E120	Embraer EMB 120	76	3.31 %
9	DC9	McDonnell-Douglas DC9	72	3.14 %
10	B73V	Boeing 737	50	2.18 %

4.2 Observations

This section presents observations made during analysis of the data, which provide detailed information about a specific aircraft flight in the CTAS study. These observations are included before the results so that the reader can better understand the methodology, and therefore better understand the statistics and data presented in Section 4.3. While each observation details a typical flight, the errors are not necessarily representative of common occurrences. Appendix C provides additional anomalous flights, which were selected to verify the methodology and to examine trajectory accuracy errors with CTAS.

4.2.1 CTAS1

This flight is a DC9 flying from Dallas/Fort Worth International Airport (DFW) to the Minneapolis-St. Paul International Airport (MSP). It departed via TEX6 through the ZEMMA intersection and proceeded to the Tulsa VORTAC (TUL). From TUL it took J25 to MSP, passing through Kansas City, Des Moines, and Mason City. The cruising altitude was 29,000 feet. The first part of the flight's filed route from DFW to ZEMMA, to TUL and past is shown in Figure 4.2-1.

4.2.1.1 Track Data

The HCS radar track started at 9,500 feet west of DFW and headed initially toward the ZEMMA intersection. About halfway there, the aircraft switched its heading toward the TUL waypoint. The horizontal track is shown in Figure 4.2-1 and in Figure 4.2-3 where the West-East scale (X axis) has been expanded by a factor of 4 to better show the location of the predicted trajectories relative to the track.

During the climb out from DFW to 29,000 feet the aircraft leveled off at 24,000 feet for three minutes before continuing the climb. The aircraft exits the ZFW airspace at level cruise at 29,000 feet. The altitude profile is shown in Figure 4.2-4.

As described in detail in Section 2.4.3, RDTRACKS processed the HCS track which included 195 position reports. First, the time intervals between track reports were examined. There were 35 of the 194 time differences between successive position reports that were equal to 11 seconds. These were changed to 12 seconds. There were 37 reports with a 13 second time difference that were changed to 12 seconds. There was one 10 second time difference that was changed to 12 seconds. There was one 14 second time difference that was changed to 12 seconds. Finally, there were two reports with a 23 second time difference that were changed to 24 seconds.

The first two reports were discarded because of inconsistent altitude values. Another track report defined as stationary had XYZ values of the immediately preceding report. The values of XYZ for this report are replaced with interpolated values. Two reports occur 24 seconds after the immediately preceding report rather than 12 seconds later. An additional interpolated track report is inserted to fill the gap in each case.

4.2.1.2 Trajectory Data

Figure 4.2-2 presents the track time line (labeled "Track") and the time line for 23 of the 168 trajectories recovered for this aircraft. Each of the trajectories is labeled with the trajectory's build time. The sample points for calculating the trajectory accuracy metrics are shown by arrows drawn from the track time line to the latest trajectory available at that sample time. The first sample starts 40 seconds after the time of the initial interpolated track point, which in this example was at 84480 seconds. 19 of the 23 trajectories shown were sampled. The aircraft departed the ZFW Center airspace at 86210 and therefore the data from the last 4 trajectories were not used.

Plots of these trajectories are shown in Figure's 4.2-1, 4.2-3, and 4.2-4. The first 6 sampled trajectories predicted the aircraft would fly to the ZEMMA intersection. After the flight flew by the ZEMMA intersection, the trajectories (Trajectory 7 and later) predicted a flight to TUL. By the eighth sampled trajectory the predicted speed and altitude matched the track.

The first five trajectories predicted the aircraft to climb to 29,000 feet; Trajectories 6 and 7 climbed the aircraft to 23,400 feet and 24,000 feet respectively. Later trajectories climbed the aircraft to 29,000 feet except for Trajectory 10 which climbed the aircraft to 28,500 feet.

4.2.1.3 Metrics

Table 4.2-1 shows the trajectory metrics calculated for this aircraft. The longitudinal and lateral errors are in nautical miles; the vertical errors are in feet. As discussed in Section 2.5.1, a sample is taken 40 seconds after the start of track and then repeated each two minutes until either the track ends, the trajectory ends, or the track leaves the center. At each sample time the distance between the track and trajectory was calculated at the current time and at look ahead times of 300 seconds or five minute increments into the future; resulting in look ahead times of 0, 300, 600, 900, 1200, 1500, and 1800 seconds.

The data shows that the lateral and longitudinal errors, although very small at low look ahead times because CTAS builds a new trajectory with each new track point, increased at the higher look ahead times early in the flight. This is because the aircraft flew inside the ZEMMA waypoint and flew direct to TUL.

It can be seen in Figure 4.2-4 that the initial estimates of climb rate were too high. By Trajectory 5 the estimate matched the actual track climbing rate. The interim altitude of 24,000 feet confuses the prediction of the final cruising altitude. Both the errors in estimating the climb rate and the errors in predicting the cruising altitude produce the large vertical prediction errors listed in Table 4.2-1.

Table 4.2-1 also shows that metrics were not computed after time 86160 because the aircraft departed the ZFW airspace at 86210.

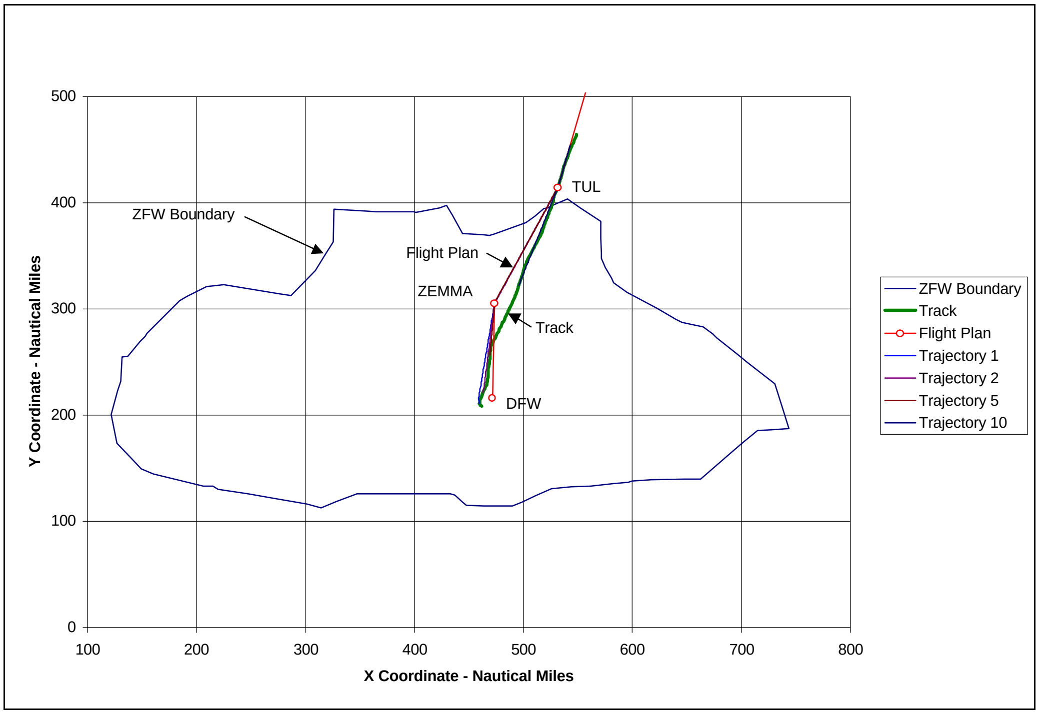

Figure 4.2-1: Aircraft Track and Route

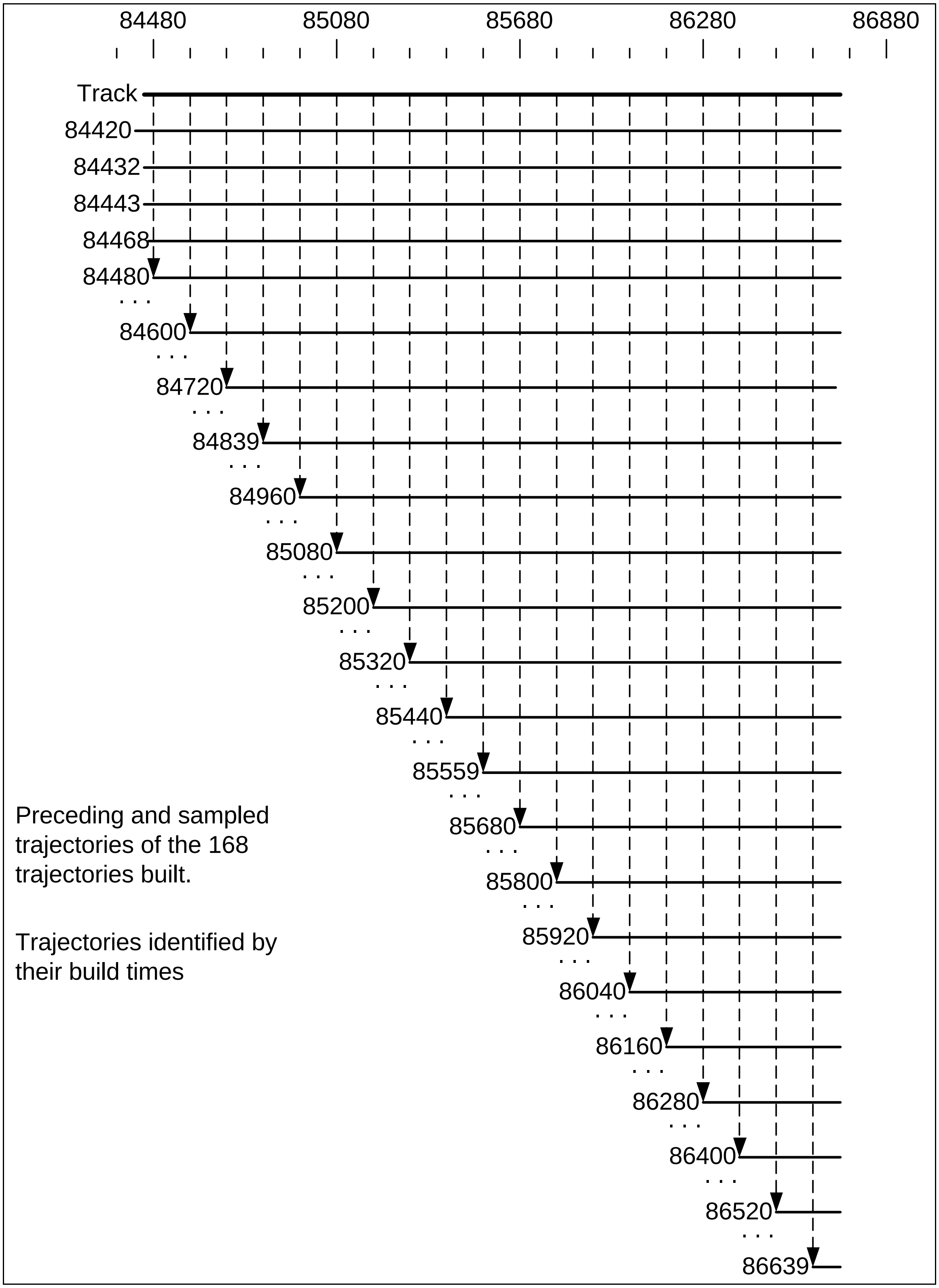

Figure 4.2-2: Sampled Trajectories

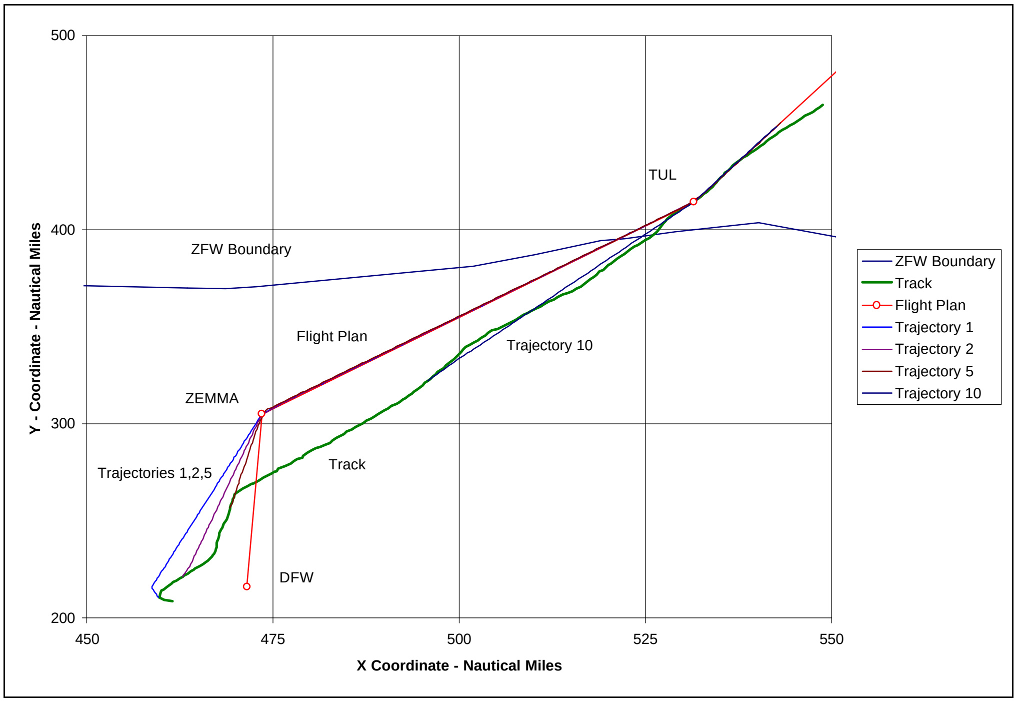

Figure 4.2-3: XY Track and Trajectories

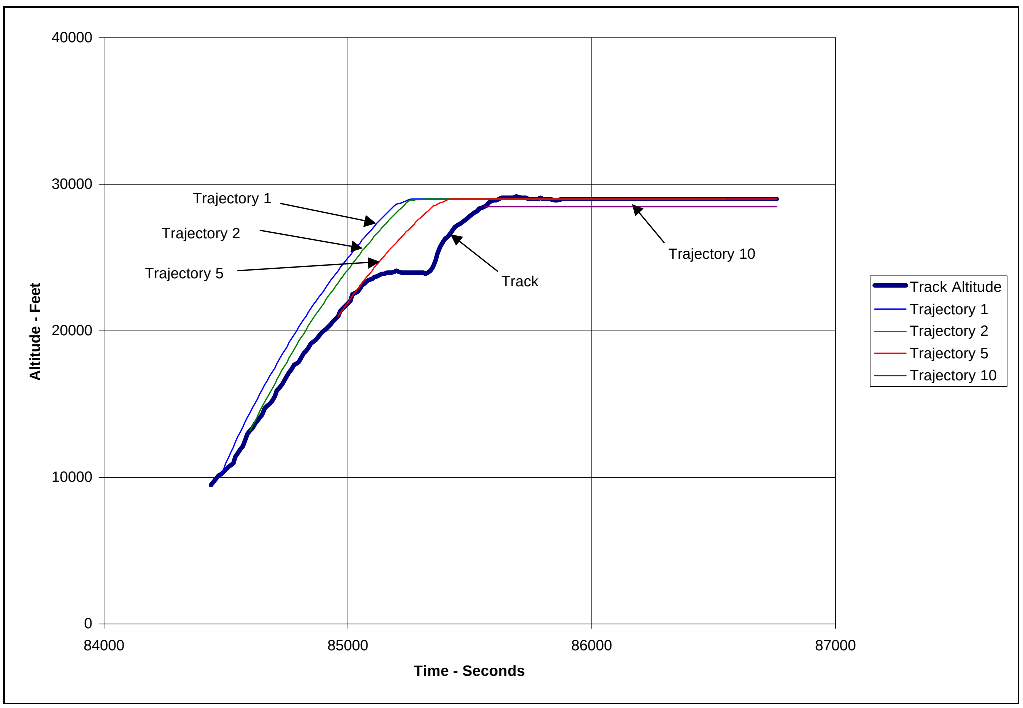

Figure 4.2-4: Altitude and Trajectory

Table 4.2-1: Trajectory Metrics (1 of 2)[9]

Sample Time	Traj No	Traj Build Time	Look Ahead Time	Long Error	Lat Error	Vert Error
84480	1	84480	0	0.00	0.01	0.00
			300	2.70	4.90	-2028.92
			600	3.84	4.77	-3188.23
			900	7.82	14.83	-3314.00
			1200	14.03	9.07	86.00
			1500	18.40	7.07	-14.00
84600	2	84600	0	0.00	0.00	0.00
			300	2.45	1.93	-1852.62
			600	4.67	8.12	-3975.92
			900	12.21	14.03	-1212.00
			1200	15.64	7.50	-12.00
			1500	19.43	4.63	-12.00
84720	3	84720	0	0.00	0.00	100.00
			300	0.48	0.07	-1088.23
			600	0.59	12.14	-4848.54
			900	5.07	10.70	-7.00
			1200	6.16	7.08	-7.00
84840	4	84839	0	-0.10	0.00	-25.74
			300	1.14	4.61	-1788.82
			600	8.53	13.93	-1912.00
			900	8.30	8.11	-12.00
			1200	11.19	5.76	-12.00
84960	5	84960	0	0.00	0.00	0.00
			300	0.77	9.81	-3105.08
			600	6.13	11.99	-502.00
			900	6.25	7.35	-102.00
			1200	7.13	3.77	-2.00
85080	6	85080	0	0.00	-0.01	0.00
			300	4.61	14.54	2300.00
			600	16.67	9.41	5700.00
			900	23.88	7.24	5600.00

[9] In this chart, longitudinal and lateral error are reported in hundredths of nautical miles, and the vertical error is reported in hundredths of feet. The precision of the input HCS altitude data is reported to the nearest 100 feet, the apparent difference is simply an artifact of the track report processing.

Table 4.2-1: Trajectory Metrics (2 of 2)

Sample Time	Traj No	Traj Build Time	Look Ahead Time	Long Error	Lat Error	Vert Error
85200	7	85200	0	0.00	0.00	100.00
			300	2.49	0.96	3800.00
			600	5.64	-0.45	5000.00
			900	10.13	1.09	5000.00
85320	8	85320	0	0.00	-0.01	-100.00
			300	-1.84	0.19	408.31
			600	-1.96	0.95	-10.00
85440	9	85440	0	0.00	-0.01	100.00
			300	-1.13	-1.42	-11.00
			600	0.64	1.07	-11.00
85560	10	85559	0	-0.11	0.00	0.00
			300	2.32	0.00	400.00
			600	5.44	0.98	500.00
85680	11	85680	0	0.00	0.00	100.00
			300	1.52	2.29	0.00
85800	12	85800	0	0.00	0.00	0.00
			300	0.80	1.19	0.00
85920	13	85920	0	0.00	0.00	0.00
86040	14	86040	0	0.00	0.00	0.00
86160	15	86160	0	0.00	0.00	0.00

4.3 Results

After running CTAS (i.e. Daisy View Release 990105) with the seven hour scenario file defined in Section 4.1, a total of 32,162 trajectories were sampled out of 352,742 trajectories. The sampled trajectories were from 2168 flights. Therefore, each one of these flights on average had 14.8 trajectories analyzed. The average duration of extracted trajectories is approximately 27 minutes with a standard deviation of nine minutes. This is lower than the actual trajectory duration built by CTAS, due to the recording process adapted in collecting these trajectories. If a trajectory exists, it is recorded at each HCS track report update (i.e. around every 12 seconds), but the actual duration recorded is only up to 32.5 minutes into the future. This is explained in more detail in Sections 2.5.1 and 4.1. The sampling process reduced the trajectory to the portion where both HCS track data and the predicted trajectory overlap in time, so the duration of the trajectory actually analyzed was reduced to approximately 22 minutes on average with a standard deviation of 11 minutes.

To set the context of the study as defined in Section 2.6.2.1, the counts of the event areas illustrated in Figure 2.6-1 are listed in Table 4.3-1 below. Referring to Figure 2.6-1, the ratio of area "a" to the sum of areas "a" and "c" defines the DST's fraction of valid flights with sampled trajectory prediction. For CTAS, 94.5 percent of the valid aircraft flights had sampled trajectory prediction.

Table 4.3-1: Valid Track and Trajectory Counts for CTAS Scenario

	Valid HCS Flight Data	Insufficient Valid HCS Flight Data	Total Flights With Trajectories
Trajectory	2168 (a)	331(b)	2499 (a +b)
Insufficient Trajectory	127 (c)		
Total Valid Flights	2295 (a + c)		

As defined in Section 2.6.2.2, another statistic useful in setting the context of the study estimates the trajectory prediction coverage over the track time analyzed. For CTAS, each analyzed flight had an average of 87 percent of prediction coverage with a standard deviation of 17.1 percent. Referring to Figure 4.3-1 and the Quantiles in Table 4.3-2, the distribution is relatively spread out with around a 99 percent of prediction coverage value at the ninetieth percentile to a 62 percent of prediction coverage value at the tenth percentile. The distribution forms a 95 percent confidence interval around the mean between 86.3 to 87.7. The maximum ratio of prediction coverage for CTAS was 99.5 percent and the minimum was 4.3 percent.

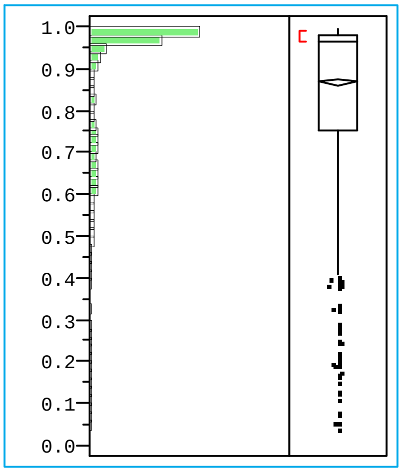

Figure 4.3-1: CTAS's Distribution of Ratio of Coverage Statistic

Table 4.3-2: Quantile Table of Ratio of Prediction Coverage

Quantile Labels	Percentiles	Values
Maximum	100.0%	0.99514
	99.5%	0.99357
	97.5%	0.99121
	90.0%	0.98780
Quartile	75.0%	0.98253
Median	50.0%	0.96952
Quartile	25.0%	0.75380
	10.0%	0.61926
	2.5%	0.45230
	0.5%	0.16663
Minimum	0.0%	0.04225

As described in Section 2.6.2.3, another descriptive value that defines the context of the analysis is the age of the trajectory at the look ahead time of zero. For CTAS, trajectories are built every time the HCS track positions are reported (every 12 seconds). There are situations where trajectories are older, including instances where CTAS did not update the trajectory or when the HCS did not supply a track exactly every 12 seconds. This study's sampled CTAS trajectories have an average trajectory age of approximately 14.6 seconds with a standard deviation of 57 seconds.

As discussed above, CTAS builds trajectories approximately every 12 seconds. The build time in seconds combined with the aircraft identifier string and HCS CID should uniquely represent a particular trajectory. However, there are instances that an aircraft has multiple trajectories with common build times. It was determined that the x and y coordinates within these multiple trajectories were close, but not identical. With the first recorded trajectory often being the correct one, the altitudes did vary significantly. Since these multiple instances occurred infrequently, it was decided to accept the first trajectory, and discard the others. Out of the 352,742 recorded trajectories in this study only 1.8 percent had more than one trajectory with a common build time.

The actual trajectory metrics and sampling process is defined in Section 2.5.1. For this seven hour ZFW scenario, 127,460 samples were taken against the 32,162 trajectories discussed above. Each sample consisted of spatial prediction error measurements including horizontal error, lateral error, longitudinal error, and vertical error. These measures are reported as a function of different look ahead times from zero to 30 minutes in the future, so the trajectory prediction performance includes the spatial prediction errors partitioned by look ahead time. As a review, look ahead time is the predicted time into the future measured from the sample start time for that particular flight. In this study increments of five minutes were used up to a look ahead time of 30 minutes into the future. In other words, if the flight had both a sampled trajectory and sufficient HCS track reports for the full range of time overlap, error measurements would be calculated at zero, five, 10, 15, 20, 25 and 30 minutes into the future for each sample at the current time.

Table 4.3-3 lists the types of statistical analyses that were performed on each of the identified factors. The analyses include either descriptive statistics in which simple tables are presented, inferential statistics in which hypothesis testing of the means and variances were performed, or both. This table also lists whether graphical information was presented with references to the appropriate section number. Inferential statistics and graphical plots (i.e. histograms and quantile tables) were calculated for a subset of the available look ahead times, including zero, 600, 1200, and 1800 seconds. The signed values of the error metrics (e.g. average lateral error) were used for these more exhaustive inferential techniques, since the sample mean acts as a measure of the bias of the trajectory predictions and the standard deviation as a measure of the uncertainty. The absolute value statistics (e.g. average absolute value of lateral error), which are also a useful measure of the uncertainty, have been included in the descriptive statistics reported in Appendix A.2.

Table 4.3-3: CTAS Analysis Summary

Factor For Samples at All Altitudes / Above FL180	Descriptive Statistics	Inferential Statistics	Histograms / Quantiles	Section Number
Look Ahead Time	Yes	Yes	Yes	4.3.1
Flight Type	Yes	Yes	No	4.3.2
Phase of Flight Horizontal	Yes	Yes	No	4.3.3
Phase of Flight Vertical	Yes	Yes	No	4.3.4

4.3.1 Analysis of Look ahead time on Trajectory Accuracy

The main factor analyzed in this study was look ahead time, defined in Section 2.2.3.3. One would expect look ahead time to have a statistically significant effect on performance, but the magnitude of the effect is also of interest. A complete table of the spatial prediction error statistics are presented at the look ahead times of zero, 300, 600, 900, 1200, 1500, and 1800 seconds (i.e. zero to 30 minutes) in Appendix A.2. The focus of the following analysis is on the signed error for lateral, longitudinal, horizontal, and vertical errors at the look ahead times of zero, 600, 1200, and 1800 seconds. This analysis includes an example set and summary results of several tables of statistical information provided by the SAS-JMP Software package (SAS Institute, 1995). They are used to evaluate the error data categorized by look ahead time and in the later sections by horizontal and vertical phase of flight. Complete tables for the CTAS data are provided in Appendix A.2. The tables present test results for unequal variance including the Levene Test and the Welch Anova Test. They also include a pairwise means comparison, referred to as the Tukey-Kramer HSD Test. Graphical plots present a comparison of means with a quantile box, a plot of the means at look ahead time versus error, and a plot of means using the Tukey-Kramer criteria.

4.3.1.1 Samples at all altitudes

The sample variance of the horizontal error from the four look ahead times are compared first by a Levene Statistical Test (Neter, 1996). Referring to Table 4.3-4, this statistical test determines if the hypothesis of equal variances can be rejected. The hypothesis can be rejected in this case, since the variances are significantly different. From Table 4.3-4, the variance of horizontal error is increasing as the look ahead time increases.

Table 4.3-4: Tests for Equal Variances and Tests for Equal Means

Tests that the Variances are Equal (Horizontal Error)[10]				
Level (seconds)	Count	Std Dev (nm)	MeanAbsDif to Mean (nm)	MeanAbsDif To Median (nm)
0	32609	0.85	0.25	0.20
600	21908	4.95	3.45	3.17
1200	12921	8.11	5.81	5.38
1800	6657	11.22	8.21	7.56
Test	F Ratio	Deg of Freedom	DF Den	Prob>F
Levene	11959.59	3	74091	0.0000
Welch Anova testing Means Equal, allowing Std's Not Equal				
	F Ratio	Deg of Freedom	DF Den	Prob>F
	10866.43	3	18479	0.0000

Next, the sample mean for each look ahead time is compared. Referring to Table 4.3-4, the Welch test is applied which compares distributions with different variances and sample sizes. It tests whether all the group means are equal. For the horizontal error at different look ahead times, the Welch Test provides evidence to reject the hypothesis that these mean errors are equal. In Figure 4.3-2, diamonds are drawn around each mean representing the 95 percent confidence interval (in this case, the diamonds are flat and look more like heavy lines due to the large range between the group means). These confidence intervals show an increase in the average horizontal error from zero to 1800 seconds look ahead time of approximately 10.6 nautical miles, from 0.3 to 10.9 nautical miles.

[10] Mean Absolute difference to mean and median are intermediate calculations in the Levene Test described in the Appendix A.0.

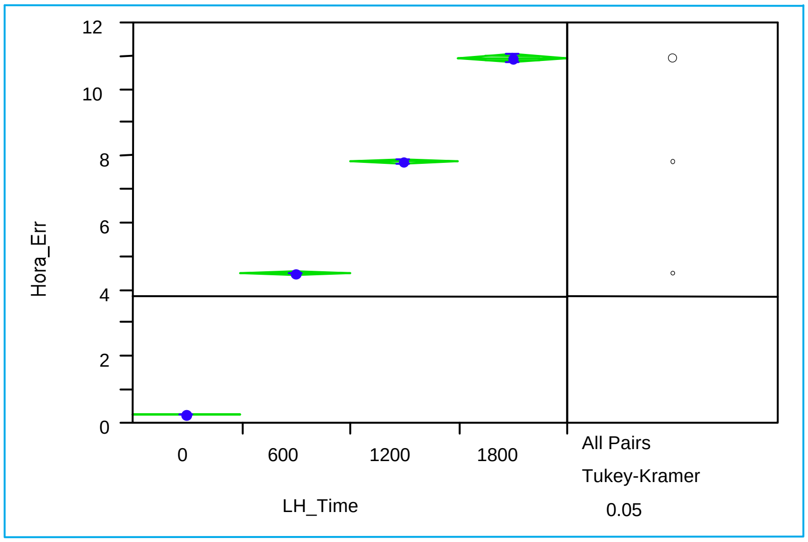

Figure 4.3-2: Sample Mean Comparison of Horizontal Error at Four Look Ahead Times[11]

The lower portion of Table 4.3-5 presents the results of a third statistical test, called the Tukey-Kramer Test, that compares all pairs of means and holds the Type I error at 0.05 for the entire test. It has the exact Type I error if the sample sizes are equal, and is conservative if they are not, which is the case in this study (Devore, 1987). The horizontal error at the four look ahead times is significantly difference between all pairs. The Tukey-Kramer Test provides a distance referred to as the Least Significant Difference (LSD)[12] that can be subtracted from the absolute difference of each pair of means. If the result is positive, the absolute difference of the means is greater than LSD, and the pair of means is significantly different. If the result is negative, the LSD is greater, and the pair is not significantly different. The upper portion of Table 4.3-5 lists the pairwise differences of the sample means for the various look ahead times. All these pairwise comparisons of the means of the horizontal error at the different look ahead times were significant.

The right side of Figure 4.3-2 presents a graphical form of the Tukey-Kramer Test. Too small to be drawn in some cases, it constructs circles around the sample means with a diameter approximately equal to the 95 percent confidence interval. However, this interval is expanded to account for the comparison of all pairs. In short, if the circles overlap the means are not considered significantly different; if they do not overlap, the means are considered significantly different. The circles drawn in Figure 4.3-2 are not overlapping at all, illustrating the numerical results that all the means are different.

[11] Normally, the height of the diamond is the length of the confidence interval and the width is proportional to the sample size. In this study, the width has been set equal for all sample sizes.
[12] LSD is proportional to the square root of the sum of the squared product of q* and the standard error of both means being compared. The q* value is a quantile similar to the t value of a Student t distribution but expanded to account for the alpha being held constant for the entire set of comparisons (SAS Institute, 1995).

Table 4.3-5: Statistical Comparison of All Means (Horizontal Error)

Means Comparisons				
Dif=Mean[i]-Mean[j]	1800	1200	600	0
1800	0.0000	3.1195	6.4127	10.6661
1200	-3.1195	0.0000	3.2932	7.5466
600	-6.4127	-3.2932	0.0000	4.3534
0	-10.6661	-7.5466	-4.3534	0.0000
Comparisons for all pairs using Tukey-Kramer HSD				
q* = 2.56909	Alpha=0.05			
Abs(Dif)-LSD	1800	1200	600	0
1800	-0.2454	2.9059	6.2146	10.4757
1200	2.9059	-0.1761	3.1361	7.3994
600	6.2146	3.1361	-01353	4.1298
0	10.4757	7.3994	4.1298	-0.1109
Positive values show pairs of means that are Significantly different.				

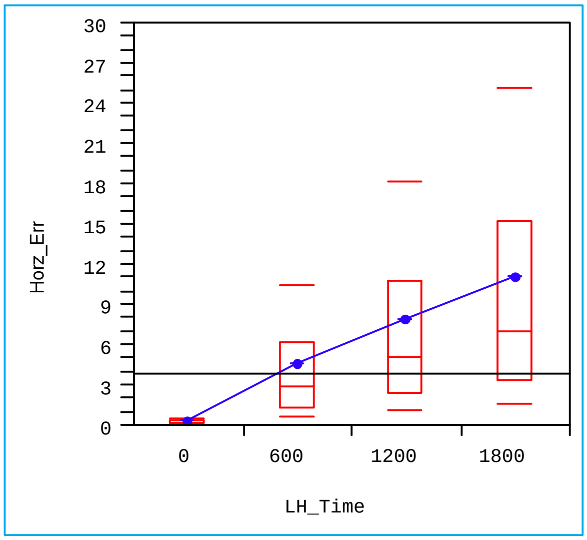

Figure 4.3-3: Quantile / Mean Comparison of Horizontal Error Vs. LH

In summary, the mean horizontal error is statistically significant at the look ahead times of zero, 600, 1200, and 1800 seconds. Referring to Figure 4.3-3, the sample means are also increasing as the look ahead time (LH) increases, ranging from a sample mean of 0.28 nautical miles at look ahead zero to 10.94 at 1800 seconds (i.e. 30 minutes). The mean of all observations is drawn as a horizontal line across the entire plot. The median is also increasing from 0.14 nautical miles at zero look ahead time to 6.9 at 1800 seconds. The horizontal lines in Figure 4.3-3's boxes correspond to the 10, 25, 50, 75, and 90 percentiles of the distribution of the sampled horizontal errors, respectively[13]. Tested statistically with the Levene Test earlier, the box ranges illustrate that the spread of the horizontal error is also increasing as the look ahead time increases.

The analysis continues by examining the lateral, longitudinal, and vertical errors using the same methods described for the horizontal error. The results are summarized in Table 4.3-6 and the means comparisons of the lateral, longitudinal and vertical errors are shown in Figures 4.3-4 through 4.3-6. The descriptive statistics of the absolute values of the four errors are tabulated in Appendix A.2.

Table 4.3-6: Statistical Results LH 0-30 minutes at All Altitudes

Error Type	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	Yes	Yes	Yes-all	Means and variance increase with look ahead time (LH).
Lateral	Yes	Yes	Yes-5of6	Only LH 1200 versus LH 1800 not different. Means (all positive) and variance increase with LH except at LH 1200 and 1800.
Long.	Yes	Yes	Yes-all	Means and variance increase with LH.
Vertical	Yes	Yes	Yes-all	Means all negative and different. Means and variance increase with LH.

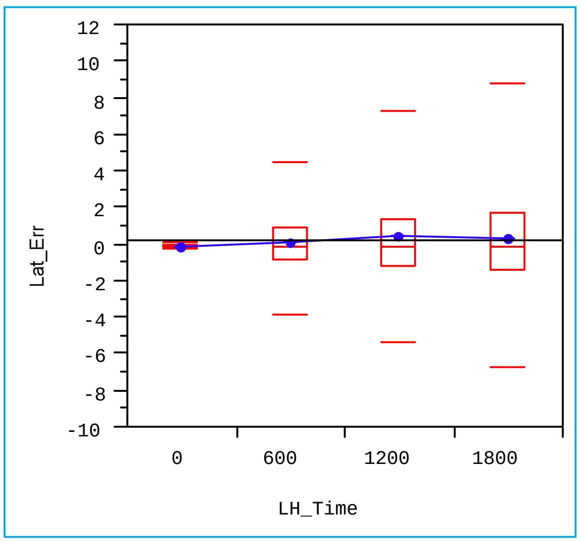

Figure 4.3-4: Quantile / Mean Comparison of Lateral Error Vs. LH

[13] The percentiles illustrated in Figure 4.3-3 as horizontal lines and box ends are described in detail in Appendix A.0.

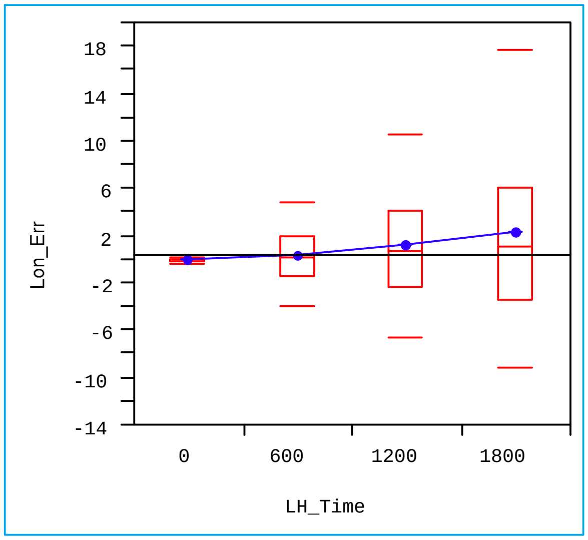

Figure 4.3-5: Quantile / Mean Comparison of Longitudinal Error Vs. LH

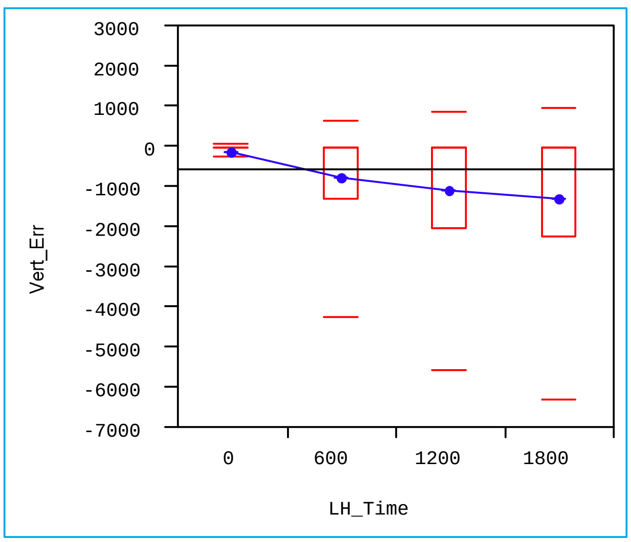

Figure 4.3-6: Quantile / Mean Comparison of Vertical Error Vs. LH

4.3.1.2 Samples at altitudes above 18,000 feet

For samples at altitudes above 18,000 feet only, the results are summarized in Table 4.3-7. The detailed histograms and statistical tables are located in Appendix A.2.

Table 4.3-7: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	Yes	Yes	Yes-all	Means and variance increase with LH.
Lateral	Yes	Yes	Yes-5of6	Only LH 1200 versus LH 1800 are not different. Variance increases with LH.
Long.	Yes	Yes	Yes-5of6	Only LH 1200 versus LH 1800 are not different. Mean and variance increases with LH.
Vertical	Yes	Yes	Yes-5of6	Means negative. Only LH 1200 versus LH 1800 are not different. LH 600 largest error. Variance increases with LH.

4.3.1.3 Discussion of the effect of look ahead time

In general, look ahead time does have a significant effect on each sample mean, which increases as the look ahead time increases. For horizontal error, the sample means increase over 10 nautical miles from zero to 1800 seconds (i.e. 30 minutes) look ahead time. The variance of the horizontal error also increases with look ahead time with a standard deviation ranging from around one nautical mile to over 11 nautical miles. Lateral and longitudinal errors are exact orthogonal components of the horizontal error, but the dominant source of horizontal error is the longitudinal error. Referring to Figures 4.3-4 and 4.3-5, the average lateral error ranges from zero to 0.46 nautical miles, and the longitudinal error ranges from slightly less than zero to around 2.4 nautical miles. The magnitude increases substantially when looking at the absolute values of the lateral and longitudinal errors. Referring to Appendix A.2, the absolute value (i.e. unsigned) means of lateral error range from 0.1 to 4.9 nautical miles from zero to 30 minutes look ahead time. The absolute value means of longitudinal error range from 0.2 to 8.1 nautical miles from zero to 30 minutes look ahead time. The vertical error mean and variance also increases for zero to 30 minutes look ahead time from –98 to –1270 feet and 790 to 3870 feet, respectively.

For the most part, the analysis of samples above 18,000 feet are consistent with the all altitudes analysis except for vertical error which seems to peak around 10 minutes (600 seconds) look ahead time at around –280 feet and actually gets less at 30 minutes to around –130 feet. The causes for this effect have been left for future analysis.

4.3.2 Analysis of Flight Type on Trajectory Accuracy

Flight type is determined by examining the origin and destination airports in a flight plan. The flight type includes four possible levels referred to as overflight, departure, arrival, and internal. Overflight is an aircraft whose origin and destination are outside the particular center's airspace, ZFW in this case. Departures leave an airport inside the center, and arrivals land at an airport inside the center. The internals include flights that have both origin and destination airports inside the center.

The analysis that follows examines whether the means of the trajectory prediction errors of the flight types are significantly different at the four look ahead times of zero, 600, 1200, and 1800 seconds. This analysis focuses on these four look ahead times and flight types against the signed lateral, longitudinal, vertical, and horizontal errors. Appendix A.2 contains a more complete set of look ahead times and also includes the descriptive statistics on the unsigned or absolute values of the errors. Figures 4.3-7 through 4.3-10 plot the sample means for each flight type as a function of look ahead time (LH) where OVR denotes overflights, ARR denotes arrivals, DEP denotes departures, and INR denotes internals.

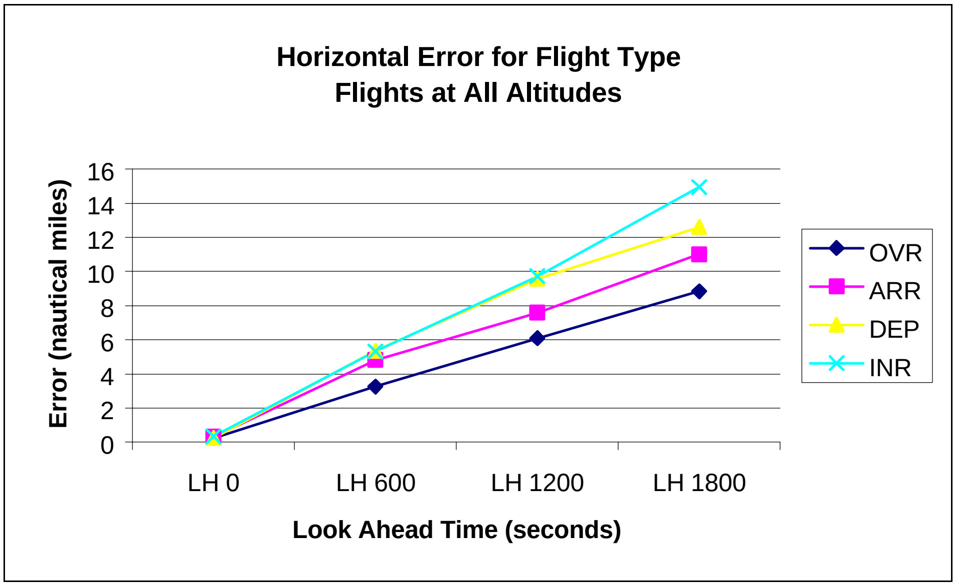

Figure 4.3-7: Sample Means for Horizontal Error per Flight Type and LH

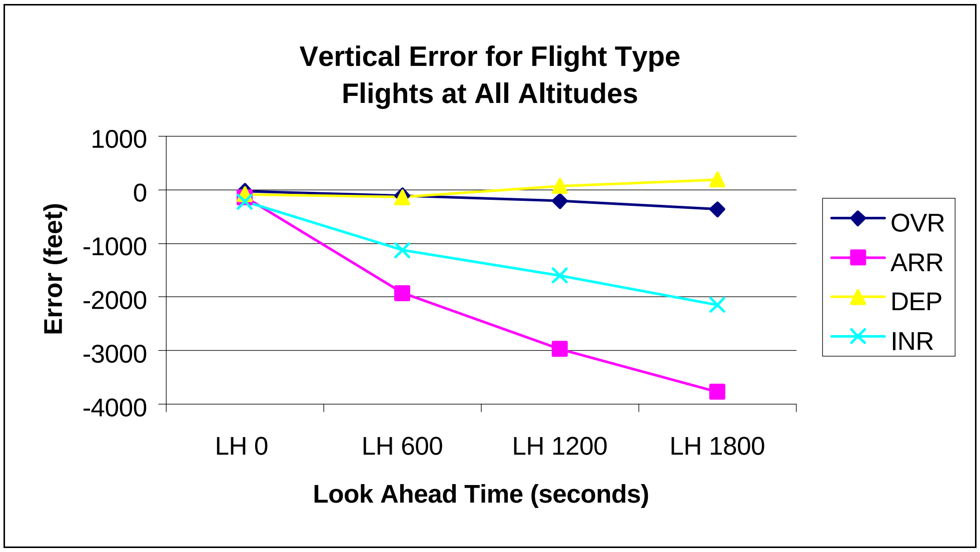

Figure 4.3-8: Sample Means for Vertical Error per Flight Type and LH

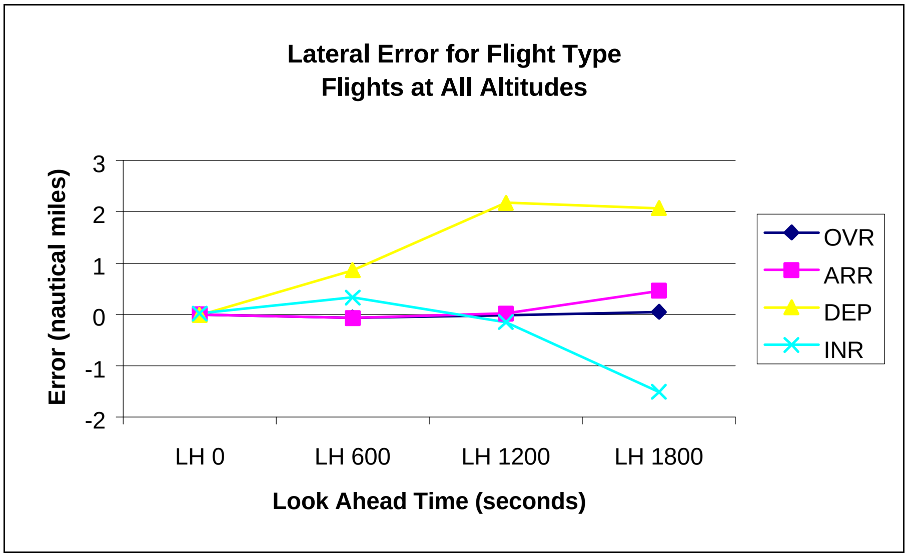

Figure 4.3-9: Sample Means for Lateral Error per Flight Type and LH

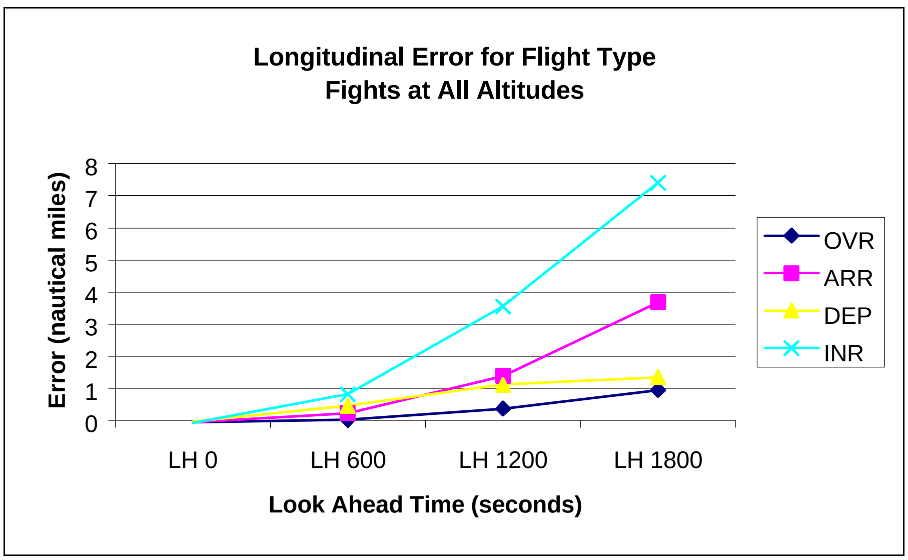

Figure 4.3-10: Sample Means for Longitudinal Error per Flight Type and LH

4.3.2.1 Samples at all altitudes

The results are summarized in Table 4.3-8. The detailed histograms and statistical tables are located in Appendix A.2.

Table 4.3-8: Statistical Results LH 0-30 minutes at All Altitudes

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-4of6	Internals versus arrivals and departures versus overflights are not different. Internals/arrivals have the largest error.
Lateral	0	Yes	Yes	Yes-1of6	Only internals versus departures significantly different.
Long.	0	Yes	Yes	Yes-1of6	Only internals versus departures different.
Vertical	0	Yes	Yes	Yes-all	All means are significantly different statistically but the magnitude is only a few hundred feet.
Horizontal	600	Yes	Yes	Yes-5of6	Internals versus departures not different.
Lateral	600	Yes	Yes	Yes-5of6	Only arrivals and overflights not different.
Long.	600	Yes	Yes	Yes-all	Maximum range between means 0.8 nm.
Vertical	600	Yes	Yes	Yes-5of6	Departures versus overflights not different. Arrivals having largest mean but internals with largest variance.
Horizontal	1200	Yes	Yes	Yes-5of6	Only internals versus departures are not different.
Lateral	1200	Yes	Yes	Yes-3of6	Only departures (with a larger error) are significantly different from the others.
Long.	1200	Yes	Yes	Yes-5of6	Only departures versus arrivals are not different. Internals have largest error.
Vertical	1200	Yes	Yes	Yes-all	Arrivals have largest error and departures smallest.
Horizontal	1800	Yes	Yes	Yes-all	Overflights have the smallest horizontal error, while internals have the largest error.
Lateral	1800	Yes	Yes	Yes-5of6	Only arrivals and overflights not different
Long.	1800	Yes	Yes	Yes-5of6	Only departures versus overflights are not different. Internals have largest error.
Vertical	1800	Yes	Yes	Yes-all	Arrivals have largest error.

4.3.2.2 Samples at altitudes above 18,000 feet

The results are summarized in Table 4.3-9. The detailed histograms and statistical tables are located in Appendix A.2.

Table 4.3-9: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-3of6	Only internals versus others are significantly different.
Lateral	0	Yes	Yes	Yes-3of6	Only internals versus others are significantly different.
Long.	0	Yes	Yes	Yes-4of6	Departures versus overflights and arrivals versus overflights are not different.
Vertical	0	Yes	Yes	Yes-3of6	Only internals versus others are different. Internals being slightly larger and positive on average while the others are negative.
Horizontal	600	Yes	Yes	Yes-all	Internals have largest error and overflights smallest.
Lateral	600	Yes	Yes	Yes-3of6	Departures (larger) different than others.
Long.	600	Yes	Yes	Yes-5of6	Only arrivals versus overflights not different. Internals have largest error.
Vertical	600	Yes	Yes	Yes-5of6	Internals versus departures are not different. Arrivals have largest error.
Horizontal	1200	Yes	Yes	Yes-5of6	Arrivals versus overflights are not different. Internals have the largest error
Lateral	1200	Yes	Yes	Yes-3of6	Departures have the largest mean and are significantly different from the others.
Long.	1200	Yes	Yes	Yes-5of6	Only overflights versus arrivals are not different. Internals have the largest mean.
Vertical	1200	Yes	Yes	Yes-all	All significantly different, but arrivals have much larger mean error and internals have much larger variance relative to the others.
Horizontal	1800	Yes	Yes	Yes-4of6	Arrivals versus overflights and departures and internals are not different. Departures and internals have the larger error.
Lateral	1800	Yes	Yes	Yes-4of6	Departures are different from others and overflights versus internals are different as well.
Long.	1800	Yes	No	No	Only variance is different, with internals having the largest variance.
Vertical	1800	Yes	Yes	Yes-5of6	Departures versus internals not different. Arrivals largest mean and internals largest variance.

4.3.2.3 Discussion of the effect of flight type

In general, flight type has a significant effect on trajectory performance. For horizontal error, overflights have the least errors as look ahead time increases, while internals have the most error ranging from 0.3 to 15 nautical miles from zero to 30 minutes look ahead time, respectively. For vertical error, arrivals seem to have the greatest mean as look ahead time increases, but internals have the largest standard deviation overall. At the lower look ahead times, the vertical error sample means vary little between flight types, but as look ahead time increases they spread out in general very quickly. For example, at look ahead time of 600 seconds or 10 minutes, the arrivals have a mean vertical error of –1923 feet while the overflights have –106 feet mean vertical error.

As far as lateral error, only departures seem to increase considerably as look ahead time increases from –0.01 to 2 nautical miles from 0 to 1800 seconds look ahead time, respectively. Longitudinal error on the other hand does increase as look ahead increases from –0.08 to 7.4 nautical miles on average. . For longitudinal error sample means, the internals dominate from around zero to 6 nautical miles larger than the other flight types on average.

4.3.3 Analysis of Horizontal Phase of Flight on Trajectory Accuracy

Horizontal phase of flight is calculated for each HCS track report and extracted for the trajectory accuracy measurements. This factor is categorized into two levels: straight or turn. The PHASE_D program that detects turns, described in Section 2.4.6.1, had its parameters set to protect against noise in the track data. As a result, rapid turns are detected but shallow turns may be missed. A turn is determined by a nine degree angle (or greater) generated by the two segments drawn from the previous position to the current position and the current position to the next position report.

The analysis that follows examines whether the mean of the trajectory prediction error at the two horizontal phases of flight are significantly different statistically at the four look ahead times of zero, 600, 1200, and 1800 seconds. This analysis will focus on these four look ahead times and two phases of flight against the signed lateral, longitudinal, vertical, and horizontal errors. Appendix A.2 contains a more complete set of look ahead times and also includes the descriptive statistics on the unsigned or absolute values of the errors. The following Figures 4.3-11 to 4.3-14 plot the sample means for each horizontal phase of flight as a function of look ahead time (LH), where STR denotes straight and TRN denotes turning.

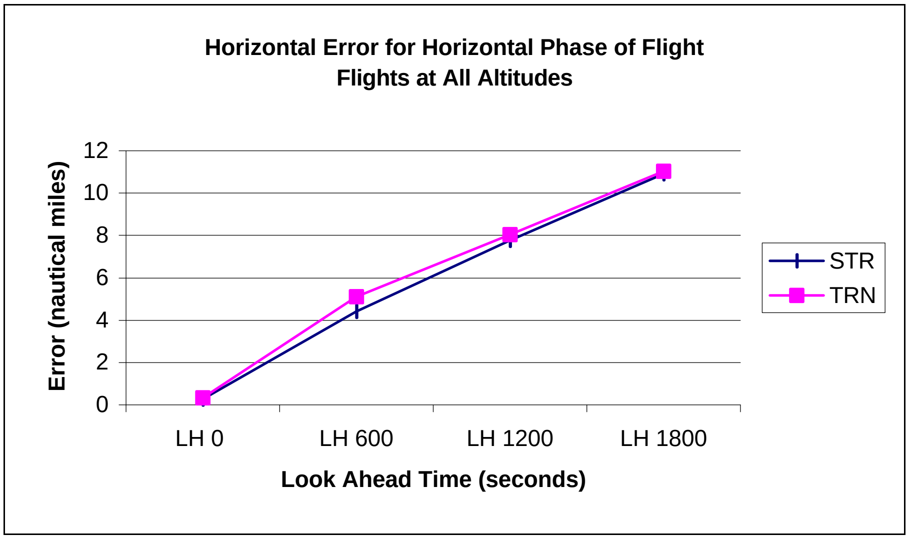

Figure 4.3-11: Sample Means for Horizontal Error per Horizontal Phase of Flight and LH

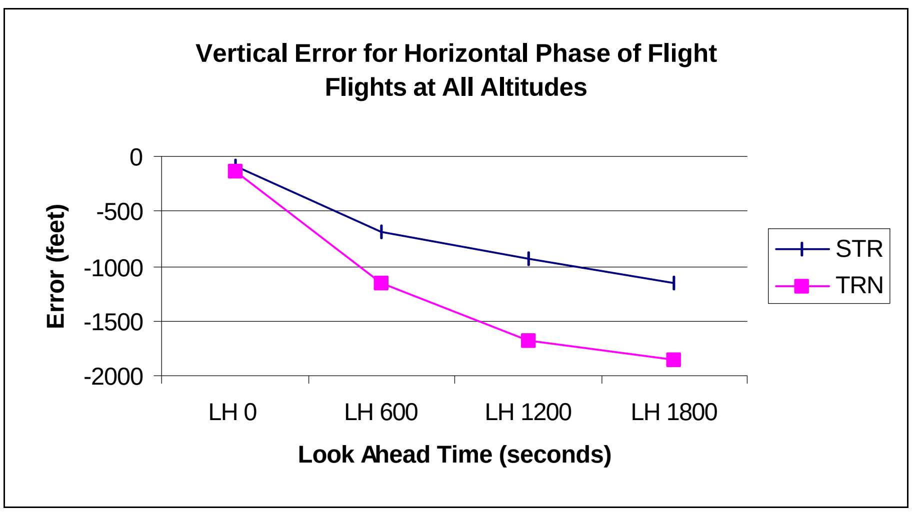

Figure 4.3-12: Sample Means for Vertical Error per Horizontal Phase of Flight and LH

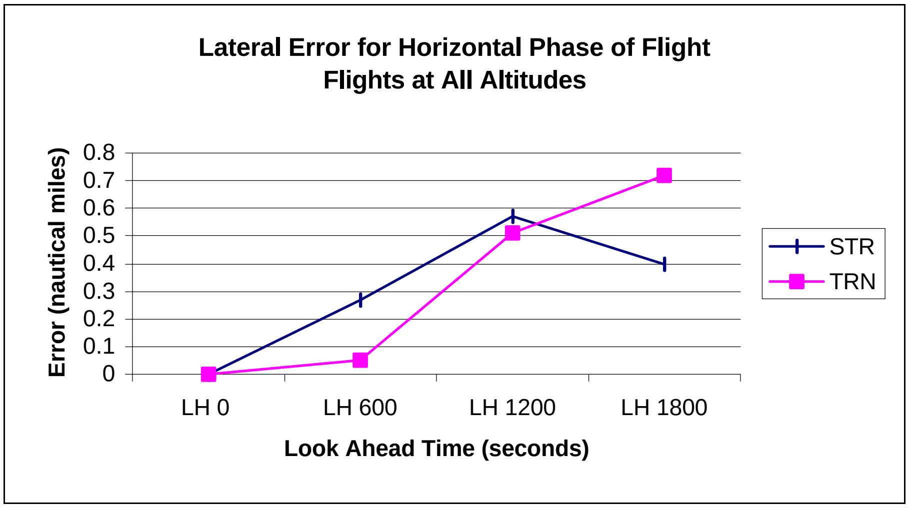

Figure 4.3-13: Sample Means for Lateral Error per Horizontal Phase of Flight and LH

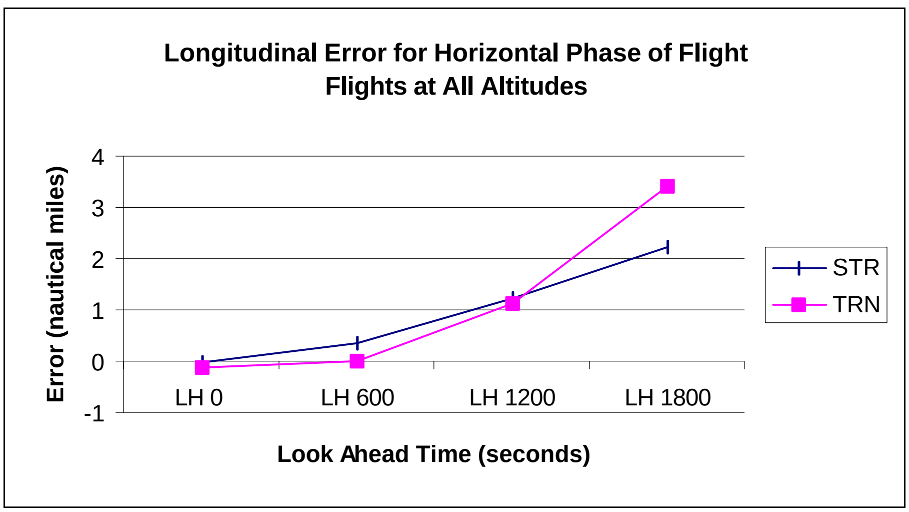

Figure 4.3-14: Sample Means for Longitudinal Error per Horizontal Phase of Flight and LH

4.3.3.1 Samples at all altitudes

The results are summarized in Table 4.3-10. The detailed histograms and statistical tables are located in Appendix A.2.

Table 4.3-10: Statistical Results LH 0-30 minutes at All Altitudes

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes	Means all different. Turns are larger by 0.07 nautical miles.
Lateral	0	Yes	No	No	Variance is different only.
Long.	0	Yes	Yes	Yes	Means both negative with turns larger by 0.07 nautical miles.
Vertical	0	Yes	Yes	Yes	Means both negative and different. Turns larger by 37 feet.
Horizontal	600	Yes	Yes	Yes	Turns larger by 0.7 nautical mile.
Lateral	600	Yes	Yes	Yes	Straight is larger by 0.22 nautical miles.
Long.	600	Yes	Yes	Yes	Straight is larger by 0.34 nautical miles.
Vertical	600	Yes	Yes	Yes	Turns larger by 460 feet.
Horizontal	1200	No	No	No	Not significantly different.
Lateral	1200	No	No	No	Not significantly different.
Long.	1200	Yes	No	No	Only variance significantly different.
Vertical	1200	Yes	Yes	Yes	Turns larger by 740 feet.
Horizontal	1800	No	No	No	Not significantly different.
Lateral	1800	No	No	No	Not significantly different.
Long.	1800	No	Yes	Yes	Turns larger around 1.2 nautical miles.
Vertical	1800	Yes	Yes	Yes	Turns larger by 700 feet.

4.3.3.2 Samples at altitudes above 18,000 feet

The results are summarized in Table 4.3-11. The detailed histograms and statistical tables are located in Appendix A.2.

Table 4.3-11: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes	Turns larger by 0.12 nautical miles.
Lateral	0	Yes	No	No	Only variance significantly different.
Long.	0	Yes	Yes	Yes	Turns larger by 0.09 nautical miles.
Vertical	0	Yes	Yes	Yes	Different but turns larger by only 30 feet.
Horizontal	600	Yes	Yes	Yes	Turns larger by 1 nautical mile.
Lateral	600	Yes	No	Yes	Only variance significantly different. T-K Test does provide evidence that means are different but Welch Test with p-value of 0.08 has more power to differentiate.
Long.	600	Yes	Yes	Yes	Turns larger by 0.23 nautical miles.
Vertical	600	Yes	Yes	Yes	Turns larger by 500 feet.
Horizontal	1200	No	No	No	Not significantly different.
Lateral	1200	No	No	No	Not significantly different.
Long.	1200	No	No	No	Not significantly different.
Vertical	1200	Yes	Yes	Yes	Turns larger by 700 feet.
Horizontal	1800	Yes	No	No	Only variance significantly different.
Lateral	1800	No	No	No	Not significantly different.
Long.	1800	No	No	No	Not significantly different.
Vertical	1800	Yes	Yes	Yes	Turns larger by 500 feet.

4.3.3.3 Discussion of the effect of Horizontal Phase of Flight

In general for horizontal error, the phase of flight in the horizontal dimension is significant only at the lower look ahead times. As the look ahead times get larger, the difference between samples at turns or straight paths becomes insignificant. However, for vertical error the difference is significant and consistently higher at all look ahead times for turns compared to straight samples. It also becomes larger as look ahead time increases. For both the horizontal and vertical dimensions, the differences between turning and straight samples is still rather small (i.e. less one nautical mile for horizontal error and 700 feet for vertical error). These small magnitudes may be caused by the insensitivity in characterizing a turn. The track points are only evaluated at large turns (around nine degrees) to protect against noise in the data, making it less powerful in detecting small turns. There has also been some discussion on the need for analysis a small distance before and after the actual turn. The current technique for determining an aircraft is turning is not sufficiently robust in filtering out the noise of the HCS track reports nor can it examine the straight path around the turn. As a result, the statistical analysis of the effect of turns should be interpreted advisedly and the algorithm will be revisited in the future.

4.3.4 Analysis of Vertical Phase of Flight on Trajectory Accuracy

Similar to horizontal phase of flight, vertical phase of flight is calculated for each interpolated HCS track report and extracted for the trajectory accuracy measurements. Vertical phase of flight is categorized into three categories: level, ascending, or descending. The track points are only labeled as climbing or descending for reasonably large climbs and descents to protect against noise in the position data, but this also prevents detection of low rate climbs and descents (i.e. smaller than 900 feet per minute). A climb or descent is determined by calculating the difference in altitude between the current interpolated track position and the next track position. If the absolute difference is less than 150 feet, the current position of the aircraft is considered in level flight, otherwise the aircraft is in a climb or descent depending on the direction up or down. Since the track positions are interpolated at 10 second intervals, the required gradient for the climbing or descending aircraft is greater than or equal to 15 feet per second or 900 feet per minute. The phase of flight algorithm is described in detail in Section 2.4.6.

The analysis that follows examines whether the mean of the trajectory prediction error at the three vertical phases of flight are significantly different statistically at the four look ahead times of zero, 600, 1200, and 1800 seconds. This analysis focuses on these four look ahead times and three phases of flight against the signed lateral, longitudinal, vertical, and horizontal errors. Appendix A.2 contains a more complete set of look ahead times and also includes the descriptive statistics on the unsigned or absolute values of the errors. The following Figures 4.3-15 to 4.3-18 plot the sample means for each vertical phase of flight as a function of look ahead time (LH), where LEV denotes level flight, ASC denotes ascending and DES denotes descending.

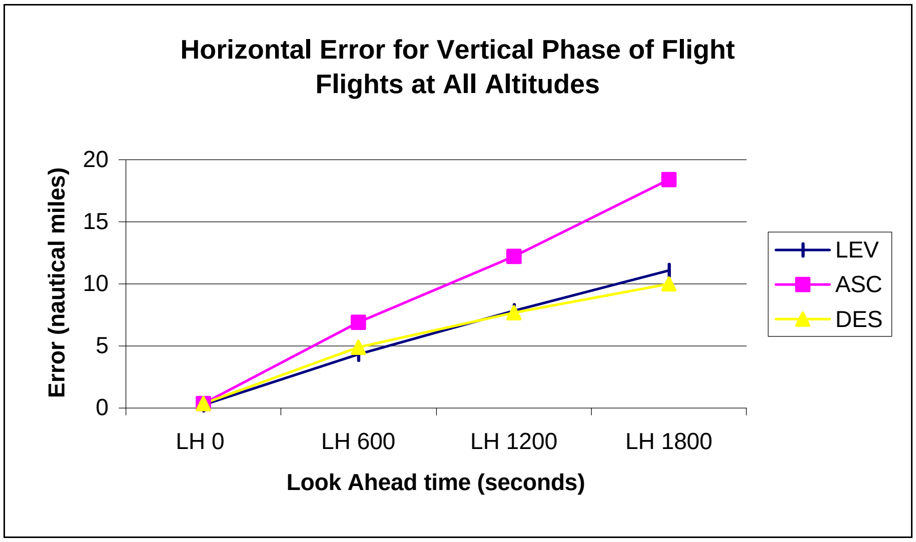

Figure 4.3-15: Sample Means for Horizontal Error per Vertical Phase of Flight and LH

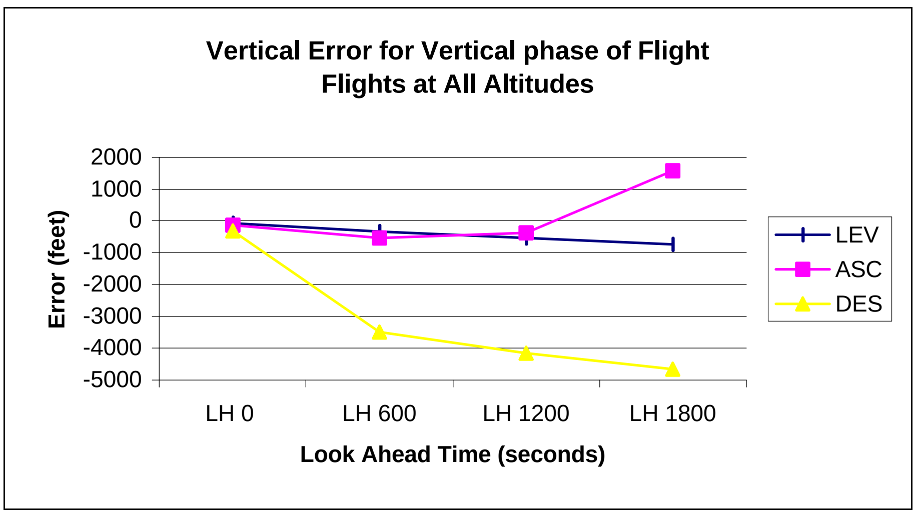

Figure 4.3-16: Sample Means for Vertical Error per Vertical Phase of Flight and LH

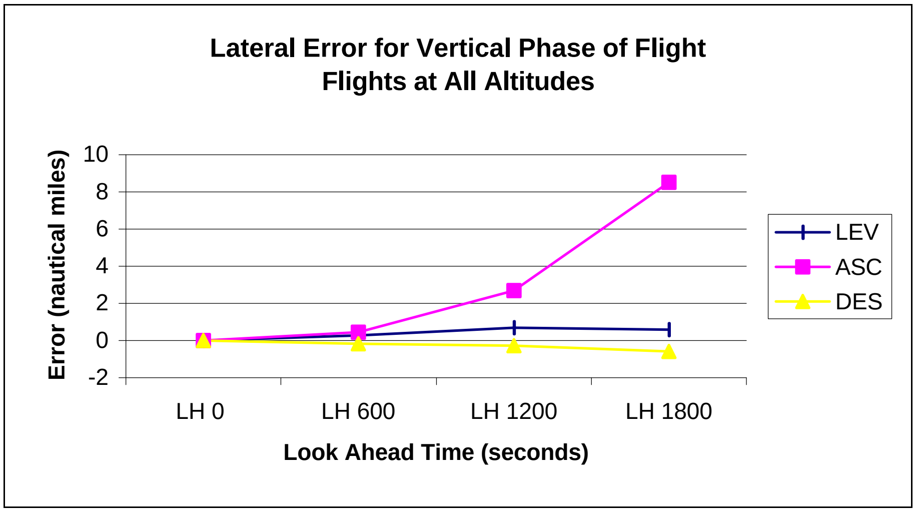

Figure 4.3-17: Sample Means for Lateral Error per Vertical Phase of Flight and LH

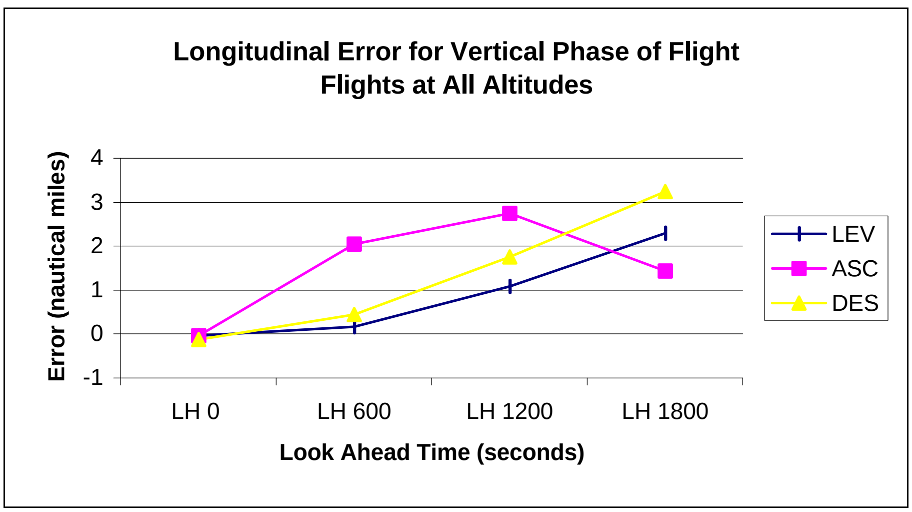

Figure 4.3-18: Sample Means for Longitudinal Error per Vertical Phase of Flight and LH

4.3.4.1 Samples at all altitudes

The results are summarized in Table 4.3-11. The detailed histograms and statistical tables are located in Appendix A.2.

Table 4.3-12: Statistical Results LH 0-30 minutes at All Altitudes

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-2of3	Level different from others. Ascent and descent same, larger error.
Lateral	0	Yes	No	No	Only variance significantly different.
Long.	0	Yes	Yes	Yes-2of3	Only ascent versus level not different.
Vertical	0	Yes	Yes	Yes	Descent largest error, -322 feet.
Horizontal	600	Yes	Yes	Yes-all	Level has largest error at 6.92 nautical miles (nm).
Lateral	600	Yes	Yes	Yes-2of3	Only ascent versus level not different.
Long.	600	Yes	Yes	Yes-all	Ascent has largest error, 2 nm.
Vertical	600	Yes	Yes	Yes-all	Descent has largest error, -3486 feet.
Horizontal	1200	Yes	Yes	Yes-2of3	Only level versus descent not different.
Lateral	1200	Yes	Yes	Yes-all	Ascent has largest error at 2.7 nm.
Long.	1200	No	Yes	Yes-1of3	Only descent versus level are different.
Vertical	1200	Yes	Yes	Yes-2of3	Only level versus ascent not different.
Horizontal	1800	Yes	Yes	Yes-all	Ascent has largest error, 18.4 miles. Inconclusive with ascent only 13 samples.
Lateral	1800	Yes	Yes	Yes-all	Ascent has largest error, 8.5 miles. Inconclusive with ascent only 13 samples.
Long.	1800	Yes	No	No	Only variance significantly different. Inconclusive with ascent only 13 samples.
Vertical	1800	Yes	Yes	Yes-2of3	Only level versus ascent not different. Inconclusive with ascent only 13 samples.

4.3.4.2 Samples at altitudes above 18,000 feet

The results are summarized in Table 4.3-12. The detailed histograms and statistical tables are located in Appendix A.2.

Table 4.3-13: Statistical Results LH 0-30 minutes Above 18,000 feet

Error Type	Look ahead Time	Levene Test	Welch Test	Tukey-Kramer	Observations
Horizontal	0	Yes	Yes	Yes-all	Ascent has largest error, 0.4 nm.
Lateral	0	Yes	No	No	Only variance significantly different.
Long.	0	Yes	Yes	Yes-2of3	Only level versus ascent not different.
Vertical	0	Yes	Yes	Yes-2of3	Only level versus ascent not different.
Horizontal	600	Yes	Yes	Yes-all	Ascent has largest error, 7 nm.
Lateral	600	Yes	Yes	Yes-2of3	Only level versus ascent not different.
Long.	600	Yes	Yes	Yes-2of3	Only level versus descent not different.
Vertical	600	Yes	Yes	Yes-all	Descent has largest error, -3033 feet.
Horizontal	1200	Yes	Yes	Yes-2of3	Only level versus descent not different. Ascent has larger error at 12.3 nm.
Lateral	1200	Yes	Yes	Yes-2of3	Only level versus ascent not different.
Long.	1200	Yes	No	No	Only variance significantly different.
Vertical	1200	Yes	Yes	Yes-2of3	Only level versus ascent not different.
Horizontal	1800	Yes	Yes	Yes-all	Ascent has largest error, 18.4 nm.
Lateral	1800	Yes	Yes	Yes-all	Ascent has largest error, 8.5 nm.
Long.	1800	Yes	No	No	Only variance significantly different.
Vertical	1800	Yes	Yes	Yes-all	Descent has largest error, -3745 feet.

4.3.4.3 Discussion of the effect of Vertical Phase of Flight

The vertical phase of flight does have a significant effect on the spatial errors. In particular, aircraft in ascent have samples with the largest horizontal mean error as look ahead time increases. From Figure 4.3-15, the sample means for ascending phase of flight range from 0.4 nautical miles to around 12 nautical miles from zero to 20 minutes look ahead time, respectively. There are only a few samples (i.e. 13 sample points) available at the larger look ahead times for ascending flight, making the results inconclusive for ascents at 30 minutes (1800 seconds) look ahead time.

The vertical phase of flight has a significant effect on vertical error as well. The descending phase of flight has the largest effect on the mean error, although the ascending samples have the largest standard deviation or variance at the lower look ahead times. Referring to Figure 4.3-16, the sample mean for descending phase of flight, which is a measure of the prediction bias, shows a decreasing (becomes more negative) average vertical error as look ahead time increases. Therefore, the trajectory prediction tends to overestimate the altitude. For aircraft in descent at look ahead times from five minutes to 30 minutes, the CTAS trajectory tends to predict either the altitude lagging (i.e. not descending fast enough), leaving the predicted altitude above the actual, or it may have lagged on its predicted location of the top of descent point, which has a similar effect.

The uncertainty of the prediction on the vertical dimension is measured by the standard deviation for each vertical phase of flight. Referring to Appendix A.2, the lower look ahead times between zero and five minutes show ascending phase of flight dominates with ranges of the standard deviation between 1400 and 4300 feet. For the larger look ahead times above five minutes, the descending phase of flight samples dominate with standard deviations ranging from 3500 to 4800 feet.

5. Summary

This report presents the results of an independent analysis of the accuracy of the trajectory modelers implemented in the URET and CTAS prototypes. These results are based on the completion of the first phase of a planned two phased effort. As originally envisioned, efforts during Phase 1 would develop a generic methodology to measure trajectory prediction accuracy which would be validated by applying it to CTAS and URET at their currently adapted sites. In Phase 2, the methodology would be applied to URET and CTAS systems that had been adapted to a common site and supplied with the same scenario. As such, the results from Phase 2 would have provided a common set of results based on the same site and scenario, allowing a comparison to be made of the two trajectory modelers, in support of research into the performance requirements for a common en route trajectory model. Unfortunately, due to funding cuts ACT-250 was only able to complete Phase 1. The results from this phase do provide the FAA with an independent set of scenario-based trajectory accuracy statistics for each DST, however, they cannot be used to compare the two DSTs due to the confounding site-specific factors.

A methodology was developed and CTAS and URET were measured based on one scenario each from their currently adapted sites (Fort Worth and Indianapolis, respectively). Both scenarios were approximately seven to 7.5 hours in duration and contained about 2500 flights. In the URET scenario from Indianapolis Center (ZID) used for this study, approximately 45 percent of the flights were overflights, 27 percent were departures, 25 percent were arrivals, and 3 percent were denoted "internals". For the CTAS scenario from Fort Worth Center (ZFW), the flight type mix was very different with approximately 13 percent of the flights being internals, 31 percent arrivals, 30 percent departures, and only 26 percent overflights. The differences in the scenarios for the flight type highlight the major differences between the scenarios and are one example why the Phase 1 results can only be reviewed individually.

The evaluation methodology took the point of view of the Air Traffic Controller using the DST. That is, a Controller viewing the aircraft predicted position data on the graphical user interface of the DST would ask how accurate the predictions were into the future, e.g., 5 minutes, 10 minutes, 20 minutes, and beyond. The Controller is not necessarily interested in the interior workings of the tool, e.g., how recently the tool made its currently valid predictions, but rather how accurate the prediction is now, and into the future. Built upon this conceptual point of view of the user, a sampling process was used to obtain the measurement data. At selected times the actual position of the aircraft was obtained from the HCS radar track data and was compared with the position of the aircraft predicted by the tool.

The Phase 1 study measured the spatial error between trajectory predictions versus the HCS track position reports, which were assumed to be the ground truth location of the aircraft. The spatial error consisted of horizontal and vertical errors. The horizontal error was further partitioned into two geometric components, lateral and longitudinal errors, representing the cross track and along track prediction errors. These errors were calculated for trajectories where both HCS track data and the DST trajectory overlapped in time. In a sense, a DST could incur higher accuracy with small trajectory errors if it selectively built trajectories; however, in this study both CTAS and URET made predictions on most of the available valid flights (aircraft movements that have both flight plan and verified track position information). For URET, 97 percent of the flights were analyzed and for CTAS 95 percent were analyzed.

The focus of the analysis was on the overall trajectory accuracy of each DST, not on individual errors. A statistical analysis was performed on the overall accuracy of the two modelers in their respective Centers with their respective scenarios. This analysis was performed on approximately 17,000 URET trajectories and 32,000 CTAS trajectories. The spatial errors have been summarized with descriptive statistics in the horizontal, lateral, longitudinal, and vertical dimensions as a function of look ahead time. Inferential statistics were performed to determine whether specific factors (i.e., look ahead time, flight type, horizontal phase of flight, and vertical phase of flight) had a significant effect on these performance statistics.

For URET, the sample means for the horizontal error, as a function of look ahead time, range from 1.2 to 10.2 nautical miles for 0 to 30 minutes look ahead time. The sample standard deviations range from 1.1 to 10.9 nautical miles. For CTAS, the sample means for the horizontal error as a function of look ahead time, range from 0.3 to 10.9 nautical miles for 0 to 30 minutes look ahead time. The sample standard deviations range from 0.9 to 11.2 nautical miles. For both URET and CTAS, the average and standard deviation of the horizontal error increases as look ahead time increases. In other words, the horizontal uncertainty of the trajectory predictions analyzed in this study increased by about 10 nautical miles on average as look ahead increased from zero to 30 minutes into the future.

As previously stated, while the Phase 1 analysis cannot be used to compare the URET and CTAS trajectory modelers, the results do provide the FAA with an independent scenario based set of trajectory accuracy measurements for each DST. All of the data from this study is stored in a large set of Oracle database tables in the WJHTC TFM Laboratory. This data can be made available to other members of the FAA community who may wish to analyze other factors, or answer other questions of interest, related to the trajectory prediction accuracy of URET and CTAS upon formal request to ACT-250. In addition, a generic methodology has been developed for the performance measurement of a common trajectory model. In FY99, this methodology and the parsing tools developed in this study will be applied to the development of DSR Workload Scenarios to be used for URET CCLD accuracy testing. With the planned adaptation of URET and CTAS to a common site, tentatively scheduled to occur in 2001, and anticipated funding availability in FY01, ACT-250 hopes to resume work on the proposed Phase 2 study to further address the FAA's efforts to determine the feasibility of a common en route trajectory model.

References

Bilimoria, K., “A Methodology for the Performance Evaluation of a Conflict Probe”, AIAA-98-4238, AIAA Guidance, Navigation and Control Conference, Boston, MA, August 1998

Brudnicki, D., *Algorithmic Evaluation Capability (AEC) Set 1 Report*, F022-L-042, MITRE/CAASD, 1995

Brudnicki, D., Arthur, W., Lindsay, K., *URET Scenario-based Functional Performance Requirements Document*, MTR98W0000044, MITRE/CAASD, 1998

Byrdsong, C. et. al., *User Request Evaluation Tool URET D3 System Testing Tools, Scenarios, and Test Results,* WN 97W00000126, MITRE/CAASD, 1997

Cale, M., Kazunas, S., Paglione, M., Ryan, Dr. H., *URET Algorithm Assessment Report,* DOT/FAA/CT-97/4, WJHTC/ACT-250, 1997

Cale, M., Paglione, M., Ryan, Dr. H., Timoteo, D., Oaks, R., *URET Conflict Prediction Accuracy Report*, DOT/FAA/CT-98/8, WJHTC/ACT-250, April 1998

Cale, M., Paglione, M., Ryan, Dr. H., Timoteo, D., Oaks, R., Summerill, S., “Application of Generic Metrics to Assess the Accuracy of Strategic Conflict Probes”, 2nd USA/Europe ATM R&D Seminar, Orlando, FL, December 1998

Devore, J., *Probability and Statistics for Engineering and the Sciences, Second Edition,* Brooks/Cole Publishing Company, 1987

Hicks, C., *Fundamental Concepts in the Design of Experiments, Fourth Edition*, Saunders College Publishing, 1993

Kelton, D., Law, A., *Simulation Modeling And Analysis, Second Edition*, McGraw-Hill, Incorporated, New York, 1991

Lindsay, K., *Xeval Users Manual*, MTR97W0000030, MITRE/CAASD, 1998

Montgomery, Douglas, C., *Introduction to Statistical Quality Control, Second Edition*, John Wiley and Sons, Inc., 1991

Neter, John, et al., *Applied Linear Regression Models, Third Edition*, Irwin, 1996

SAS Institute, *JMP Statistics and Graphics Guide, Version 3, JMP Software Package*, 1995

WJHTC/ACT-250, *Generic Metrics and Statistics to Estimate the Conflict Prediction Accuracy of Conflict Probe Tools,* 1997

WJHTC/ACT-250, *Generic Metrics and Statistics to Estimate the Accuracy of Trajectory Modelers,* 1998

WJHTC/ACT-250, *Conflict Probe Data Reduction and Tools Interface Control Document*, 1999

INTENTIONALLY BLANK

List of Acronyms

ACID	Aircraft Identification
ACT-250	WJHTC ATM Engineering, Research and Evaluation Branch
ARR	Arrival
ARTCC	Air Route Traffic Control Center
ASC	Ascending
ATC	Air Traffic Control
ATM	Air Traffic Management
BOD	Bottom of Descent
CAASD	Center for Advanced Aviation System Development
CCLD	Core Capability Limited Deployment
CID	Computer Identification
CLT	Central Limit Theorem
CPP	CTAS Parser Program
CTAS	Center-TRACON Automation System
DEP	Departure
DES	Descending
DST	Decision Support Tool
ENR	En Route
FAA	Federal Aviation Administration
FL	Flight Level
FFP1	Free Flight Phase 1
FP	Flight Plan
GIM	General Purpose Output Interface Module
HCS	Host Computer System
HSD	Honestly Significant Difference
IAIPT	Interagency ATM Integrated Product Team
IFR	Instrument Flight Rules
INR	Internal
JRPD	Joint Research Project Description
LEV	Level flight
LH	Look ahead time
LSD	Least Significant Difference
MTR	Monitor Test and Recording
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
nm	Nautical Mile
OVR	Overflight
RHCMP	Reverse Host Converge/Merge Process
SAS	Statistical Analysis Systems
SID	Standard Instrument Departure
ZQL	Standard Query Language
STAR	Standard Arrival Route
STD	Standard Deviation
STR	Straight
TFM	Traffic Flow Management
TJS	Trajectory Sampling
TOD	Top of Descent
TRN	Turning

URET	User Request Evaluation Tool
WJHTC	William J. Hughes Technical Center
ZFW	Fort Worth ARTCC
ZID	Indianapolis ARTCC
ZKC	Kansas City ARTCC
ZOB	Cleveland ARTCC